

QY 930 AATGATGGCTGCATGAGC-----TCTACCTGATCCCTGGCTGAACCTGGCTGCT 983
 Db 343 AsnaBpMetLeuAlaValAsnGlnGlnAlaAsnLeuIleProGlyLeuAsnLeuSerAla 362
 QY 984 GAGAGCTTTTCCGAGCTTCATCCAGCGGAGTCCCGCCGCTT-----CCAGCAGCGCTT 1037
 Db 363 LeuGlyIlePheSerThrGlyLeuSerValLeuProProAlaGlyProArgGlyVal 382
 QY 1038 ACTGGGCGCTGCTCCCTATAGCTCCCTTATG----- 1067
 Db 383 ProProSerProProTyrHisProPheAlaThrHisSerGlyTyrPheSerSerLeuTyr 402
 QY 1068 -----CAGGCTCCCGAGCAGAGATGCTG 1091
 Db 403 ProHisHisIspHeglyProPheProHisHisHisSerTyrProGlnGlnIleThrVal 422
 QY 1092 CAGGCTTTATCCCGCCGAGCAGTGGCGCATCATCGGACAGAGAGGAGAGCATC 1151
 Db 423 SerLeuPheIleProThrGlnAlaValGlyAlaIleIleGlyIleGlyValAlaHisIle 442
 QY 1152 AAACAGCTCTCCCGGTTGGCAGCGCTTCATCAAGATTGACACACCGGAAACACCTGAC 1211
 Db 443 LysGlnLeuAlaAlaGlyPheAlaGlyAlaSerIleLysIleAlaProAlaGlyGlyProAsp 462
 QY 1212 TCCAAGTTCGTATGTTATCATCATCTGACCGCCAGAGGCCCAATTCAGAGCTCAAGGGA 1271
 Db 463 ValSerGlnArgMetValIleIleIleThrGlyProProGlnAlaGlnPheLysValSerAla 482
 QY 1272 AGA 1274
 Db 483 Lys 483

RESULT 15
 ID Q9VZ69 PRELIMINARY; PRT; 566 AA.
 AC Q9VZ69; Q9VZ69; (TREMBlrel, 13, Created)
 DT 01-MAY-2000 (TREMBlrel, 22, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel, 25, Last annotation update)
 DE CG1691 protein (IGF-II mRNA-binding protein) (SD07045p).
 OS IMP OR CG1691.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBL_TaxID=7227;
 RX NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Branton R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballow R.M., Basu R.M., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A.A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Flossler C., Gabrielle A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glisick A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodera C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Skupski M.P., Smith T.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.C., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Branton R.C., Rogers Y.,
 RA Barzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Gang N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Idegawa C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nimco J.,
 RA Paclet J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirska R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Miara S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.B., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richer J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutnack F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Nielsen J., Nielsen F.C., Christiansen J.;
 RT "Cloning and Expression of a Drosophila Homolog of IMP/Vg1-RBP.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nimco J., Paclet J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yin C., Lewis S.E., Rubin G.M., Celniker S.;
 CC Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AE003484; AAF47958.2; -;
 DR EMBL; AF241237; AAF63331.1; -;
 DR EMBL; AY069821; AAL39966.1; -;
 DR FlyBase; FBgn0030235; Imp.
 DR GO; GO:0003676; Functional acid binding; IEA.
 DR InterPro; IPR004087; KH dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH; 4.

DR SMART, SM00322, KH, 4.
DR PROSITE; PSS0084; KH TYPE 1; 4.
SQ SEQUENCE 566 AA; 62129 MW; B4C74C5C9B0C830 CRC64;

Alignment Scores:

Pred. No.:	Length:	Matches:
Score:	2.67e-55	566
Percent Similarity:	863.50	203
Best Local Similarity:	56.60%	Conservative: 93
Query Match:	38.81%	Mismatches: 157
	24.45%	Indels: 70
		Gaps: 12

US-09-270-437d-7 (1-1946) x Q9VZ69 (1-566)

```

QY 213 AACACCGAGAGTGAACCGCAGTGTGAATGTCACTTATTCACCGGAGACGACGAG 272
DB 4 AsnAsnAsnSerSerArgLeuAsnAsnAsnIleSer---AsnAsnTrpTyrGlnGlnLys 22
QY 273 CAAGCCATCATGAG-----CTGAATGGCCACCAGTTGGAG 308
DB 23 GlnSerLeuIleArgTyrLeuAspArgAlaAlaValGlyLeuAsnGlyValGlnPheGlu 42
QY 309 AACCATGCCCTGAAGTCTCTCATCCCGCATGACAGATGACACAGGAGCTGAGAT 368
DB 43 GlySerIlySleuHlaIalGlnGlnLeuAspIlysaSngln----- 55
QY 369 GGGCGCCGAGGGGGCTTGGCTCTCGGGGTCAAGCCCGCAGGGGCTCACTTGGCAGG 428
DB 56 -----ArgArgSerGlnArgAsnGlnArgAsnProTyr---Pro 67
QY 429 GGGGGCCCGACGACGAGCAGACATGCCCTTGGGCTCTGGTGGCCACCCAG 488
DB 68 GlyMetProGlyProGlyArgGlnAlaAspPheProLeuArgIleLeuValGlnSerGlu 87
QY 489 TATGTGGGTGCATTATTTGCAAGGAGGGGCCACCATCCGCAACATCAAAACAGACC 548
DB 88 MetValGlyAlaIleIleGlyArgGlnGlySerThrIleArgThrIleThrGlnGlnSer 107
QY 549 CAGTCCAGATGACGCTGATAGAAAGAGAACGACGAGTGCAGCTGAAAAAGCCATCAGT 608
DB 108 ArgAlaArgValAspValHisArgIlysgIlyAsnValGlySerLeuGlnLysSerIleThr 127
QY 609 GTGCACTCCACCCCTGAGGGGCTCTCCGCTTGAAGATGATCTTGAAGATTATGACAT 668
DB 128 IleTyrGlyAsnProGlnAsnGlySerThrAsnAlaCysIlyArgIleLeuGlnValMetGln 147
QY 669 AAAGAGGCTAAGGACCAAAACGGGCTGACAGGTTCCCTGAAGATCCTGGCCCAAT 728
DB 148 GlnGlnAlaIleSerThrAsnLysGly---GlnIleCysLeuLysIleLeuAlaHisAsn 166
QY 729 AACCTTGAAGGGCTCATTTGCAAGAAAGACGAACTGAAAGAGTGAAGCAAGAT 788
DB 167 AsnLeuIleGlyArgIleIleGlyLysSerGlyAsnThrIleLysArgIleMetGlnAsp 186
QY 789 ACCGAGACAAATATCCATCCATCTCTGTCGAAGACCTTACCTTACACCTGAGAG 848
DB 187 ThrAspThrIlySleThrValSerSerIleAsnAspIleAsnSerPheAsnLeuGlnArg 206
QY 849 ACCATCACTGTGAAGGGGCGCATCGAAGATTGTTCAGAGGCGGACGAGAAATATGAAG 908
DB 207 IleIleThrValIlysgIlyLeuIleGlyAsnMetSerArgIleGlnAsnGlnIleSerThr 226
QY 909 AAAGTTCGGAGGCTATGAGAAATGATGTGGCTGCCATGAGC-----TTCACCTGATC 962
DB 227 LysLeuArgGlnSerTyrGlnAsnAspLeuGlnAlaMetAlaProGlnSerLeuMetPhe 246
QY 963 CCGTGGCTGAACCTGGCTGCTGATAGTCTTTTCCAGACTTCA-----TTCAGC 1010
DB 247 ProGlyLeuHisProMetAlaMetMetSerThrProGlyAsnGlyMetValPheAsnThr 266
QY 1011 GCAGTCCCGCGGCT-----CCGACGAGGTTACTGGGGCTGCTCCCTAT 1055
DB 267 SerMetProPheProSerCysGlnSerPheAlaMetSerLysThrProAlaSerValVal 286

```

```

QY 1056 AGCTCTTATGACAGGCTCCGACGACGAGATGTGACAGTGTATTCCTCCCGCCAGGCA 1115
DB 287 ProProValPheProAsnAspLeuGlnGlnThrThrTyrLeuTyrIleProAsnAsnAla 306
QY 1116 GTGGGCGCCATCATTCGCGACAGAGGGCGACACATCAAAACAGCTCTCCGGGTTGGCCAG 1175
DB 307 ValGlyAlaIleIleIleGlyThrArgIlySerHisIleArgSerIleMetArgPheSerAsn 326
QY 1176 GCTTCACATCAAGATTGACACCCGAAACA-----CTGATCCCAAGTTGCT 1223
DB 327 AlaSerLeuLysIleAlaProLeuAspAlaAspIlyProLeuAspGlnGlnThrGlnArg 346
QY 1224 ATGTTATCATCACTGACCCCGCAGAGGCCCAATTCAAGGCTCAGGGAGAAATCTATAGC 1283
DB 347 LysValThrIleValGlyThrProGlnGlyGlnTrpLysAlaGlnTyrMetIlePheGlu 366
QY 1284 AACTCAAGAGAGAACTTTGTTGGTCCCAAGAGAGAAAGTGAAGTGAAGACCAACAATA 1343
DB 367 LysMetArgGlnGlnGlyPheMetCysGlyThrAspAspValArgLeuThrValGlnLeu 386
QY 1344 CGTGGCCAGCATCAGCAGCTGGCCGGTCAATTGGCAAGGTGAAAGAAACGGTGAACAG 1403
DB 387 LeuValAlaSerSerGlnValGlyArgIleIleGlyLysGlyGlyGlnAsnValArgGln 406
QY 1404 TTGCAGAAATTTGACGCGACGCTGAGGTGTAATCAACA----- 1442
DB 407 LeuGlnArgValAlaThrGlySerValIleLysLeuProGlnHisAlaLeuAlaProProSer 426
QY 1443 -----GACCAACCCCTGATGAGAAACGACGAGTCATGTAATAATCATCGACAT 1493
DB 427 GlyIlyAspGlnGlnThrPro-----ValHisIleIleGlyLeu 439
QY 1494 TTCTATGCCAGTCAAGTGCCTCAACGAAAGATCCGACATCCTGGCCAG----- 1544
DB 440 PheTyrSerValGlnSerAlaGlnArgArgIleArgAlaMetMetLeuSerThrAsnPro 459
QY 1545 -----GTTAAGCAGACATCAAGAAAGGACAGAGTAAC 1577
DB 460 ProProIleThrIlySleGlnLysAlaAlaLysGlnGlnGlnGlnGlnGlnGlnSer 479
QY 1578 CAGGCCCGACGACGAGAGAGTGAAGACGCCCTCCCTGCTTGAAGTCCAGACACAGA 1637
DB 480 LeuAlaGlyAla-AlaSerSerGlySerGlnGlnGlnGlnProGlnSerProSerGlnGln 499
QY 1638 ACGGCGCA 1644
DB 499 nglnAla 501

```

Search completed: July 23, 2004, 11:15:00
Job time : 120.039 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 03:38:17 ; Search time 12641.7 Seconds

(without alignments)
11256.058 Million cell updates/sec

Title: US-09-270-437D-8

Perfect score: 3283

Sequence: 1 ggcagcgagagagcgagga.....aaccttgaaatgttattc 3283

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_ro.*
10: gb_ro.*
11: gb_ro.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hcg_hum.*
31: em_hcg_inv.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rod.*
36: em_hcg_mam.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hcgo_hum.*
40: em_hcgo_mus.*
41: em_hcgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3282	100.0	3283	6	AR171867
2	3282	100.0	3283	6	AR343076
3	3282	100.0	3283	6	BD209927
4	3253.8	99.1	3667	9	AF057352
5	3143	95.7	3412	6	AR171865
6	3143	95.7	3412	6	AR343074
7	3143	95.7	3412	6	BD209925
8	3115	94.9	3633	9	BC021290
9	2799.4	85.3	142971	9	AC020629
10	2781.4	84.7	91084	9	AF596177
11	2779.8	84.7	182695	2	AC015706
12	2327	70.9	187226	9	AC104980
13	2325.4	70.8	113201	9	AP004290
14	1924.2	58.6	3557	10	BC023758
15	1711.8	52.1	2010	9	AF117107
16	1627	49.6	98945	9	AC016961
17	1618	49.3	2640	9	HSM806243
18	1613	49.1	208273	2	AC134942
19	993.6	30.3	4953	10	BC054552
20	748.6	22.8	156398	2	AC094924
21	734.8	22.4	2021	5	AF026527
22	718.2	21.9	2105	10	AF541940
23	707.2	21.5	50320	2	AC126373
24	700.6	21.3	2130	9	AF117106
25	697.4	21.2	2381	9	AF198254
26	697.2	21.2	1740	6	BD275902
27	697.2	21.2	1740	6	AR220687
28	697.2	21.2	1740	6	AR281251
29	697.2	21.2	1740	6	AX365954
30	697.2	21.2	1743	6	AX368054
31	697.2	21.2	1743	6	AX366057
32	695.6	21.2	4155	9	HSU76705
33	695.6	21.2	4159	6	AR171863
34	695.6	21.2	4159	6	AR343072
35	695.6	21.2	4159	6	BD209923
36	695.6	21.2	4171	9	AF117108
37	695.6	21.2	4181	6	BD275767
38	695.6	21.2	4181	6	AR220552
39	695.6	21.2	4181	6	AR255546
40	695.6	21.2	4181	6	AR281116
41	695.6	21.2	4181	6	AX333233
42	695.6	21.2	4181	6	AX365782
43	695.6	21.2	4181	9	HSU97188
44	695.6	21.2	4602	6	AX397963
45	695	21.2	1740	10	AB046173

ALIGNMENTS

RESULT 1	AR171867	3283 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	AR171867				
DEFINITION	Sequence 8 from patent US 6297364.				
ACCESSION	AR171867				
VERSION	AR171867.1				
KEYWORDS	GI:17910817				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 3283)				
AUTHORS	Chem,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K. and Old,L.J.				
TITLE	Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof				

Pred. No. is the number of results predicted by chance to have a

QY 2041 GCGTATCCCTTTTAACTGTAACATAGTGAAGCTGTTCAAGCCAAAGCAAAATGCAC 2100
Db 2041 GCGTATCCCTTTTAACTGTAACATAGTGAAGCTGTTCAAGCCAAAGCAAAATGCAC 2100
QY 2101 ACCCTTTTCTGTGGCAATTCGTCTCTGTACATGTGTGTACATTAAGAAAGGAAGATG 2160
Db 2101 ACCCTTTTCTGTGGCAATTCGTCTCTGTACATGTGTGTACATTAAGAAAGGAAGATG 2160
QY 2161 TTAAGATATGTGGCTGTGGGTACACAGGGGCTGTGACGGGTATATATTTTGAAT 2220
Db 2161 TTAAGATATGTGGCTGTGGGTACACAGGGGCTGTGACGGGTATATATTTTGAAT 2220
QY 2221 AATATATCAAAATTACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2280
Db 2221 AATATATCAAAATTACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2280
QY 2281 TTAAGAGAAAGCAGGCTTTTCTGTACCTTAAAGATTAAGCTTTTGGAGGTCTCAGG 2340
Db 2281 TTAAGAGAAAGCAGGCTTTTCTGTACCTTAAAGATTAAGCTTTTGGAGGTCTCAGG 2340
QY 2341 TGTAGAGAGAGCTTTGAGGCGACCGGCAAAATTCACCCAGAGGAAATCTGTGGA 2400
Db 2341 TGTAGAGAGAGCTTTGAGGCGACCGGCAAAATTCACCCAGAGGAAATCTGTGGA 2400
QY 2401 AGGACACTCAAGGAGTCTGGATCAGCTGTGTATGTCAACAGAGGAGTACCGTCTCT 2460
Db 2401 AGGACACTCAAGGAGTCTGGATCAGCTGTGTATGTCAACAGAGGAGTACCGTCTCT 2460
QY 2461 TGAAGAGAAACTCTGTCACTCTCAGCTCTGCTAGCTCAACACCATTTCTCTTGC 2520
Db 2461 TGAAGAGAAACTCTGTCACTCTCAGCTCTGCTAGCTCAACACCATTTCTCTTGC 2520
QY 2521 TTACAGGTTTAACTGTTTTTGTGATCTGCTATATATATTTCTGTCTCTCTGTT 2580
Db 2521 TTACAGGTTTAACTGTTTTTGTGATCTGCTATATATATTTCTGTCTCTCTGTT 2580
QY 2581 TATCTCTCCCTCCCTCCCTCCCTCCCTCTCTCATCTCATTTTGAATTTCTCAT 2640
Db 2581 TATCTCTCCCTCCCTCCCTCCCTCCCTCTCTCATCTCATTTTGAATTTCTCAT 2640
QY 2641 CCTTCCATCTCAATCCCGTATCTAGCAGCCCCCCCCCGGCAAGCAAGTCTCTGA 2700
Db 2641 CCTTCCATCTCAATCCCGTATCTAGCAGCCCCCCCCCGGCAAGCAAGTCTCTGA 2700
QY 2701 GATACATCAACCAAAAGGAACAAAGGAAACACCAACAGCTCACTTACCTT 2760
Db 2701 GATACATCAACCAAAAGGAACAAAGGAAACACCAACAGCTCACTTACCTT 2760
QY 2761 GATTACTCAAAAGACAGATCAATGTACTTGTCTAGCTTTTGGAGAGGAAACA 2820
Db 2761 GATTACTCAAAAGACAGATCAATGTACTTGTCTAGCTTTTGGAGAGGAAACA 2820
QY 2821 GGAACCAACCAACCAACCAATCAACCAACCAAGAAATTTCCAAATGAAAGATGT 2880
Db 2821 GGAACCAACCAACCAACCAATCAACCAACCAAGAAATTTCCAAATGAAAGATGT 2880
QY 2881 AATTGTCTTTTGTGATTTGTGTATAGCCATCAATATTCAGCAAAATGATTCCTTTC 2940
Db 2881 AATTGTCTTTTGTGATTTGTGTATAGCCATCAATATTCAGCAAAATGATTCCTTTC 2940
QY 2941 TTTAAAAAATATGTGAGGAAATAGAAATTTTCCAAAGTGTGTGGCCAGGGCGTT 3000
Db 2941 TTTAAAAAATATGTGAGGAAATAGAAATTTTCCAAAGTGTGTGGCCAGGGCGTT 3000
QY 3001 AATATCAAGATTTTAAACGAGAAACACACAGAGAGTACCTCAGGTGTTTTTA 3060
Db 3001 AATATCAAGATTTTAAACGAGAAACACACAGAGAGTACCTCAGGTGTTTTTA 3060
QY 3061 CCTCAGACCTTGTCTGT 3120
Db 3061 CCTCAGACCTTGTCTGT 3120

QY 3121 TTTTATTTTTTTTAAATAAATAGTGGAAAAAATTAAGATATCACTGCCAGCTTG 3180
Db 3121 TTTTATTTTTTTTAAATAAATAGTGGAAAAAATTAAGATATCACTGCCAGCTTG 3180
QY 3181 GAGAAGTGAAGTCCAAAGTGTGCAACAGCTGTTCTGAATGTCTTCGGCTAGCCAAAGA 3240
Db 3181 GAGAAGTGAAGTCCAAAGTGTGCAACAGCTGTTCTGAATGTCTTCGGCTAGCCAAAGA 3240
QY 3241 CCNATATGCGCTTCTTTTGAACAAACCTTGAATAATGTTATTT 3283
Db 3241 CCNATATGCGCTTCTTTTGAACAAACCTTGAATAATGTTATTT 3283

RESULT 2
AR343076 3283 bp DNA linear PAT 17-AUG-2003
LOCUS AR343076
DEFINITION Sequence 8 from patent US 6576756.
ACCESSION AR343076
VERSION AR343076.1 GI:33738478
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 3283)
AUTHORS Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
and Old,L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof
JOURNAL Patent: US 6576756-A 8 10-JUN-2003;
FEATURES
source
1. .3283
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 3282; DB 6; Length 3283;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGCGGAGAGAGCGAGAGCGCCGGGTACCGGGCCGGGGAGCGCGGGCTCTGGGG 60
Db 1 GGCAGCGGAGAGAGCGAGAGCGCCGGGTACCGGGCCGGGGAGCGCGGGCTCTGGGG 60
QY 61 AAGAGAGGATGATGAACAAAGCTTTTATCATGAGGAACTGAGCCCGGCTGACCGCGAC 120
Db 61 AAGAGAGGATGATGAACAAAGCTTTTATCATGAGGAACTGAGCCCGGCTGACCGCGAC 120
QY 121 GACCTCCGGAGCTTTTGGGGACAGAAAGCTGCCCTGGCGGACAGAGTCTGTGAAG 180
Db 121 GACCTCCGGAGCTTTTGGGGACAGAAAGCTGCCCTGGCGGACAGAGTCTGTGAAG 180
QY 181 TCCGGCTAGCCTTGGTGAAGTACCCCGACAGAACTGGGCAATCCGCGCATCGAGAC 240
Db 181 TCCGGCTAGCCTTGGTGAAGTACCCCGACAGAACTGGGCAATCCGCGCATCGAGAC 240
QY 241 CTCTCGGTTAAGTGAATTCATGAGGAAATATGAGGAAATGATTAATCAGTCTGTAA 300
Db 241 CTCTCGGTTAAGTGAATTCATGAGGAAATATGAGGAAATGATTAATCAGTCTGTAA 300
QY 301 AAGCTAAGAGAGAGAAATTCAGATTCGAATCATCCCTCTCACTGCAAGTGGAGGTG 360
Db 301 AAGCTAAGAGAGAGAAATTCAGATTCGAATCATCCCTCTCACTGCAAGTGGAGGTG 360
QY 361 TTGATGAGCTTTTGGCTCAATATGAGGACAGTGAAGTGGAAACAGTCAACAGAC 420
Db 361 TTGATGAGCTTTTGGCTCAATATGAGGACAGTGAAGTGGAAACAGTCAACAGAC 420
QY 421 AAGAAACCGCGCTTGTCAACGTCAATATGCAACAGAAAGAAACAAATAGCCATG 480
Db 421 AAGAAACCGCGCTTGTCAACGTCAATATGCAACAGAAAGAAACAAATAGCCATG 480
QY 481 GAGAGCTAAGGGGAGTCAAGTTGAGAACTACTCTTCAAGTTTCTTACATCCGGAT 540
Db 481 GAGAGCTAAGGGGAGTCAAGTTGAGAACTACTCTTCAAGTTTCTTACATCCGGAT 540

Db 481 GAGAAAGCTAAGCGGGCAATCAAGTTGAAGAACTTCTTCAAGATTTCCTACATCCGGAT 540
QY 541 GAAAGAGTGAAGCTCCCTTCGCCCCCTCAGAGAGCCGAGCGTGGGAGCACTTTCCTCCGG 600
Db 541 GAAAGAGTGAAGCTCCCTTCGCCCCCTCAGAGAGCCGAGCGTGGGAGCACTTTCCTCCGG 600
QY 601 GAGCAAGGCGCAGCGCCCTCGGGGGCACTTCTCAAGGCGAGACAGATTGATTTCCGCTGGG 660
Db 601 GAGCAAGGCGCAGCGCCCTCGGGGGCACTTCTCAAGGCGAGACAGATTGATTTCCGCTGGG 660
QY 661 ATCCGCGCCCCCAGTCTTGTGTGTCATCATCGGAAAGAGGGCTTGACCAATAAG 720
Db 661 ATCCGCGCCCCCAGTCTTGTGTGTCATCATCGGAAAGAGGGCTTGACCAATAAG 720
QY 721 AACATCATTAAGCAGACCCAGTCCCGGGTAGATATCATTAAGAAAGAGAACTTGGAGCT 780
Db 721 AACATCATTAAGCAGACCCAGTCCCGGGTAGATATCATTAAGAAAGAGGGCTTGACCAATAAG 780
QY 781 GCGAGGAAGCCTGTACCATCATATGCCACCCGAGGGGACTTCTGAAGCATGCCGATG 840
Db 781 GCGAGGAAGCCTGTACCATCATATGCCACCCGAGGGGACTTCTGAAGCATGCCGATG 840
QY 841 ATTCTTGAATCATCAGAAAGAGGAGCAGATGAGACCAACTAGCCGAGAGATTCCTCG 900
Db 841 ATTCTTGAATCATCAGAAAGAGGAGCAGATGAGACCAACTAGCCGAGAGATTCCTCG 900
QY 901 AAAATCTTGGCACAACATGGCTTGGTTGGAGACTGATTGGAAAGAGGCGAAATTTG 960
Db 901 AAAATCTTGGCACAACATGGCTTGGTTGGAGACTGATTGGAAAGAGGCGAAATTTG 960
QY 961 AAGAAATTTGAACATGAAACAGGAGCCAGATPAACATCTCATCTTGGAGATTGAGC 1020
Db 961 AAGAAATTTGAACATGAAACAGGAGCCAGATPAACATCTCATCTTGGAGATTGAGC 1020
QY 1021 AATATCAACCCCGAAAGAACATCACTGTGAAGGGCAGATTGAGGCTGTGCCAGTGT 1080
Db 1021 AATATCAACCCCGAAAGAACATCACTGTGAAGGGCAGATTGAGGCTGTGCCAGTGT 1080
QY 1081 GAGATGAGATTATGAAGAGCTGGCTGAGGCTTTGAAAATGATGCTGGCTGTAAAC 1140
Db 1081 GAGATGAGATTATGAAGAGCTGGCTGAGGCTTTGAAAATGATGCTGGCTGTAAAC 1140
QY 1141 ACCCACTCCGGAATCTTCCAGGCTGTACCCCATCACAGTTTGGCCGTTCCGCAAT 1200
Db 1141 ACCCACTCCGGAATCTTCCAGGCTGTACCCCATCACAGTTTGGCCGTTCCGCAAT 1200
QY 1141 ACCCACTCCGGAATCTTCCAGGCTGTACCCCATCACAGTTTGGCCGTTCCGCAAT 1200
Db 1141 ACCCACTCCGGAATCTTCCAGGCTGTACCCCATCACAGTTTGGCCGTTCCGCAAT 1200
QY 1201 CATCACTCTTATCCAGAGCAGAGATTGTGAATCTTTCATCCCAACCGAGCTGTGGC 1260
Db 1201 CATCACTCTTATCCAGAGCAGAGATTGTGAATCTTTCATCCCAACCGAGCTGTGGC 1260
QY 1261 GCCATCATCGGGAAGAGGGGGGCAACATCAACAGCTGGCGAGATTGCGCGAGCTCT 1320
Db 1261 GCCATCATCGGGAAGAGGGGGGCAACATCAACAGCTGGCGAGATTGCGCGAGCTCT 1320
QY 1321 ATTAAGATTTGCTCCGGAAGGGCCAGAGCTCAGAGGAAGAGATGTCATCATCACGGG 1380
Db 1321 ATTAAGATTTGCTCCGGAAGGGCCAGAGCTCAGAGGAAGAGATGTCATCATCACGGG 1380
QY 1381 CCACCGGAAGCCAGATTCAAGGCGCCAGAGCGGATCTTTGGGAAACTGAAAGAGGAAAC 1440
Db 1381 CCACCGGAAGCCAGATTCAAGGCGCCAGAGCGGATCTTTGGGAAACTGAAAGAGGAAAC 1440
QY 1441 TTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGATATCAAGTGGCCCTCTTCCACA 1500
Db 1441 TTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGATATCAAGTGGCCCTCTTCCACA 1500
QY 1501 GCTGGCGGGGTGATTGGCAAGGTGCAAGACCGTGAACGAATCGCAGAACTTAAACAGT 1560
Db 1501 GCTGGCGGGGTGATTGGCAAGGTGCAAGACCGTGAACGAATCGCAGAACTTAAACAGT 1560
QY 1561 GCAGAAATCATGCTGCTGTGACCAAAAGCCAGATGAAATGGAAGTGTGTGAGA 1620
Db 1561 GCAGAAATCATGCTGCTGTGACCAAAAGCCAGATGAAATGGAAGTGTGTGAGA 1620

QY 1621 AATATCGGCACTTCTTGTAGAGCCAGATGCAACGGCAAGATCAGGAAATTGTACAA 1680
Db 1621 AATATCGGCACTTCTTGTAGAGCCAGATGCAACGGCAAGATCAGGAAATTGTACAA 1680
QY 1681 CAGGTGAAGCAGAGAGCAGAAATACCTTCAAGGAGTGGCTTCAAGGCGAGAGAGTA 1740
Db 1681 CAGGTGAAGCAGAGAGCAGAAATACCTTCAAGGAGTGGCTTCAAGGCGAGAGAGTA 1740
QY 1741 GGCCTCCACAGGCGCACGCAAAACAGGATGAATGTACCTTCAACACTGACAGAA 1800
Db 1741 GGCCTCCACAGGCGCACGCAAAACAGGATGAATGTACCTTCAACACTGACAGAA 1800
QY 1801 TGAGACCAAGCAGCAGCAGCAGATCGGAGCAAAACCAAGACCATCTGAGAAATGAGAA 1860
Db 1801 TGAGACCAAGCAGCAGCAGCAGATCGGAGCAAAACCAAGACCATCTGAGAAATGAGAA 1860
QY 1861 GTCTCGGAGGCGCGCAGGAGCTGTGCGAGAGCCCTGAGAAACCCGAGGGGCGAGAGGG 1920
Db 1861 GTCTCGGAGGCGCGCAGGAGCTGTGCGAGAGCCCTGAGAAACCCGAGGGGCGAGAGGG 1920
QY 1921 GCGGGGAAGGTCAAGCAGTTTGCAGAACCAAGCCGCGCTCCGCGCCCGCCAGGAGC 1980
Db 1921 GCGGGGAAGGTCAAGCAGTTTGCAGAACCAAGCCGCGCTCCGCGCCCGCCAGGAGC 1980
QY 1981 TTCTGCAAGCTTCAAGCCATCCACTTACATTCAGTCTGATCTTCTGAACTCCACGA 2040
Db 1981 TTCTGCAAGCTTCAAGCCATCCACTTACATTCAGTCTGATCTTCTGAACTCCACGA 2040
QY 2041 CGCTATCCCTTTTATGTTGAATCAATATGAGAGAGTGTGCAAAACCAAGAAATGAC 2100
Db 2041 CGCTATCCCTTTTATGTTGAATCAATATGAGAGAGTGTGCAAAACCAAGAAATGAC 2100
QY 2101 ACCCTTTTCTGTCGCAATGCTCTGTGACATGTGTGATCATTTAGAAAGGAAAGATG 2160
Db 2101 ACCCTTTTCTGTCGCAATGCTCTGTGACATGTGTGATCATTTAGAAAGGAAAGATG 2160
QY 2161 TTAAGATATGTGGCTGTGGTTAACAAGGGTGTGCAAGGGGTAAATATTTAGAAAT 2220
Db 2161 TTAAGATATGTGGCTGTGGTTAACAAGGGTGTGCAAGGGGTAAATATTTAGAAAT 2220
QY 2221 AATATATCAAAATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 2280
Db 2221 AATATATCAAAATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 2280
QY 2281 TTAAGAGAAAGCAGGCTTTCTAGACTTTAAGATAAAGTCTTGGAGGTCTACGG 2340
Db 2281 TTAAGAGAAAGCAGGCTTTCTAGACTTTAAGATAAAGTCTTGGAGGTCTACGG 2340
QY 2341 TGTAGAGAGAGCTTGAAGGCAACCGCACAAAATTCACCCAGAGGAAATCTGTCGGA 2400
Db 2341 TGTAGAGAGAGCTTGAAGGCAACCGCACAAAATTCACCCAGAGGAAATCTGTCGGA 2400
QY 2401 AGGACACTCAGCGCAGTTCGATCACTGTGTATGTCAACAGAGGAGATACCGTCTCT 2460
Db 2401 AGGACACTCAGCGCAGTTCGATCACTGTGTATGTCAACAGAGGAGATACCGTCTCT 2460
QY 2461 TGAAGAGAAACTCTGCACTCTCATGTCTGTATGCTATACCACTTCTCTTGGC 2520
Db 2461 TGAAGAGAAACTCTGCACTCTCATGTCTGTATGCTATACCACTTCTCTTGGC 2520
QY 2521 TTCAAGGTTTAACTGTTTGTGATATGTGTATATATTTCTGTCTCTCTGTT 2580
Db 2521 TTCAAGGTTTAACTGTTTGTGATATGTGTATATATTTCTGTCTCTCTGTT 2580
QY 2581 TATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCAT 2640
Db 2581 TATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCAT 2640
QY 2641 CCGTCCATCTCAATCCGTAATCAAGCAACCCCGCCCGCCAGGCAAGCAGTCTCTGA 2700
Db 2641 CCGTCCATCTCAATCCGTAATCAAGCAACCCCGCCCGCCAGGCAAGCAGTCTCTGA 2700

Qy	2701	GATGCACTACACAAAGAAAGAACAAAGGAAAACAACAACCGAGCTTCAACTTACCTT	2766
Db	2701	GTATGCACATCACAAAAAGAAACAAAAGGAAACACAAAAACAGCTTCATTACCTT	2766
Qy	2761	GGTTACTCAAAGAACAAGATCATAGTACTTGCTTAGAGCTTTTGGAAAGGAAACA	28220
Db	2761	GGTTACTCAAAGAACAAGATCATAGTACTTGCTTAGAGCTTTTGGAAAGGAAACA	28220
Qy	2821	GGAAACCACAAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCA	28800
Db	2821	GGAAACCACAAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCA	28800
Qy	2881	AATTTGCTTTTTGCTATTTGGTGTATAAAGCCATCATATTTCAGCAAAATGATTCCTTC	29400
Db	2881	AATTTGCTTTTTGCTATTTGGTGTATAAAGCCATCATATTTCAGCAAAATGATTCCTTC	29400
Qy	2941	TTTAATAAAAAAAAAATGTGGAGGAAGTAGAAAATTACCAAGTGTGTGGCCCCAGGCGTT	30000
Db	2941	TTTAATAAAAAAAAAATGTGGAGGAAGTAGAAAATTACCAAGTGTGTGGCCCCAGGCGTT	30000
Qy	3001	AAATTACAGATTTTATTTTACGAGAAAAACACACAGAGAAGCTACCTCAGGTGTTTTTA	30600
Db	3001	AAATTACAGATTTTATTTTACGAGAAAAACACACAGAGAAGCTACCTCAGGTGTTTTTA	30600
Qy	3061	CCTCAGCACTCTGCTCTGTGTTTCCCTTAAGATTTGTAAAGCTGATAGTTGAGCAAT	31200
Db	3061	CCTCAGCACTCTGCTCTGTGTTTCCCTTAAGATTTGTAAAGCTGATAGTTGAGCAAT	31200
Qy	3121	TTTTTATTTTATTTTATTTTAAATTAAGCTGGAAAAAATAAGATATCACTGCCAGCCTG	31800
Db	3121	TTTTTATTTTATTTTATTTTAAATTAAGCTGGAAAAAATAAGATATCACTGCCAGCCTG	31800
Qy	3181	GAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATGCTTCGCGTAGCCAGAA	32400
Db	3181	GAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATGCTTCGCGTAGCCAGAA	32400
Qy	3241	CCNATATGCGCTTCTTTTGGACAAACCTTGAAAAATGTTATTT 3283	
Db	3241	CCNATATGCGCTTCTTTTGGACAAACCTTGAAAAATGTTATTT 3283	
RESULT 3			
LOCUS	BD209927	3283 bp DNA linear PAT 17-JUL-2003	
DEFINITION		Isolated nucleic acid molecules encoding cancer-associated antigens, these antigens and method of using the same.	
ACCESSION	BD209927		
VERSION	BD209927.1	GI:33019697	
KEYWORDS	JP 2002512049-A/6.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Chen,X.T., Gure,A., Teang,S., Stockert,E., Jager,E., Knuth,A. and Old,L.J.		
TITLE	Isolated nucleic acid molecules encoding cancer-associated antigens, these antigens and method of using the same		
JOURNAL	Patent: JP 2002512049-A 6 23-APR-2002;		
COMMENT	LUDWIG INSTITUTE FOR CANCER RESEARCH OS Homo sapiens (human) PN JP 2002512049-A/6 PD 23-APR-2002 PR 16-MAR-1999 JP 2000545030 PF 17-APR-1998 US 09/061709 PI YAO TSENG CHEN,ALI GURE,SOLAM TSANG,ELISABETH STOCKERT,ELKE PI JAGER, PI ALEXANDER KNUTH,LLOYD J OLD PC C12N1/5/09,A6K3/5/12,A6K3/9/00,A6K3/3/9,A6K3/5/00,C07K16/32. PC C12N1/15, PC C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/574,		

CC	C12N15/00, C12N15/00	n-not determined(3243)
PH	Key	Location/Qualifiers
FT	source	1..3283
FEATURES		/organism='Homo sapiens' (human)'. source
ORIGIN		1..3283 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"
Query Match	100.0%; Score 3282; DB 6; Length 3283;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 3283;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GGCAGCGGAGGAGGAGGAGCGGGTACCGGGCGGGGAGCCGGCGGCTCTCGGGG 60
Db	1	GGCAGCGGAGGAGGAGGAGGAGCGGGTACCGGGCGGGGAGCCGGCGGCTCTCGGGG 60
QY	61	AAGAGACGGATGATGAAACAGCTTTACATCGGGAACCTGAGCCCCGCGCTACCGCCGAC 120
Db	61	AAGAGACGGATGATGAAACAGCTTTACATCGGGAACCTGAGCCCCGCGCTACCGCCGAC 120
QY	121	GACCTCCGGCAGCTCTTTGGGGACAGGAAGCTGCCCTGGCGGACAGGTCTGCTGAAG 180
Db	121	GACCTCCGGCAGCTCTTTGGGGACAGGAAGCTGCCCTGGCGGACAGGTCTGCTGAAG 180
QY	181	TCCGCGTACGCGCTTGATGGACTACCCCGACCGAAGCTGGGCGATCGCGCGCATGTGAGACC 240
Db	181	TCCGCGTACGCGCTTGATGGACTACCCCGACCGAAGCTGGGCGATCGCGCGCATGTGAGACC 240
QY	241	CTCTCGGGTAAAGTGAATTGCGATGGGAAATCATGGAAGTTGATTACTCAGTCTTAA 300
Db	241	CTCTCGGGTAAAGTGAATTGCGATGGGAAATCATGGAAGTTGATTACTCAGTCTTAA 300
QY	301	AAGCTAAGGACAGGAAATTCAGATTGAAACATCCCTCTCTCACCTGAGTGGAGGTG 360
Db	301	AAGCTAAGGACAGGAAATTCAGATTGAAACATCCCTCTCTCACCTGAGTGGAGGTG 360
QY	361	TTGATGGACTTTGGCTCAATATGGGACAGTGGAAATGTGAAACAAGTCAACAGAC 420
Db	361	TTGATGGACTTTGGCTCAATATGGGACAGTGGAAATGTGAAACAAGTCAACAGAC 420
QY	421	ACAGAAACCGCGCTGTCTCAACGTCACTATGCAACAAGAAAGGAAATATGCGCATG 480
Db	421	ACAGAAACCGCGCTGTCTCAACGTCACTATGCAACAAGAAAGGAAATATGCGCATG 480
QY	481	GAGGAAGCTAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTACATCCGGAT 540
Db	481	GAGGAAGCTAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTACATCCGGAT 540
QY	541	GAAAGAGTGAAGTCCCTTGCGCCCTCAGCGAGCCAGGCTGGGACCACTCTTCCGG 600
Db	541	GAAAGAGTGAAGTCCCTTGCGCCCTCAGCGAGCCAGGCTGGGACCACTCTTCCGG 600
QY	601	GAGCAAGGCCACGCGCTGGGGGCACTTCTCAGGCCACAGATTAATTTCCGCTGCGG 660
Db	601	GAGCAAGGCCACGCGCTGGGGGCACTTCTCAGGCCACAGATTAATTTCCGCTGCGG 660
QY	661	ATCTGTGATCCCAACCGAGTTTGTGGTGGCCATCATGGAAGGAGGGCTTGAACATTAAG 720
Db	661	ATCTGTGATCCCAACCGAGTTTGTGGTGGCCATCATGGAAGGAGGGCTTGAACATTAAG 720
QY	721	AACATCTAAGACAGACCCAGTCCCGGATGATATCTCATTAAGAAAGAACTCTTGAGCT 780
Db	721	AACATCTAAGACAGACCCAGTCCCGGATGATATCTCATTAAGAAAGAACTCTTGAGCT 780
QY	781	GCAAGAAAGCTGTCAACATCATATGCAACCCACAGAGGGACCTTGAAGATGCGCGCATG 840
Db	781	GCAAGAAAGCTGTCAACATCATATGCAACCCACAGAGGGACCTTGAAGATGCGCGCATG 840
QY	841	ATTCTTGAATCATGCAAAAGAGGCAATGAGACCAAACTAGCCGAAAGATTCCTCTG 900

Db	841	ATTCTTGAATCATGCGAAGAAAGGACGAGATGAGACCAAACTGAGCCGAAAGATTCCTCTGG	900
Qy	901	AAAATCTTGGCACAATAGCTGCTGGTTGGAAAGCTGATATGGAAAAAAGAGCAAAATTTGG	960
Db	901	AAAATCTTGGCACAATAGCTGCTGGTTGGAAAGCTGATATGGAAAAAAGAGCAAAATTTGG	960
Qy	961	AAGAAAATTTGAACATGAAACAGGGACCAAGATTAACAATCTCATCTTTGGCAGGATTTGAGC	1020
Db	961	AAGAAAATTTGAACATGAAACAGGGACCAAGATTAACAATCTCATCTTTGGCAGGATTTGAGC	1020
Qy	1021	ATTATCAACCCGGAAAGAACCATTCATCTGTAAAGGGCACGTTGAGGCTGTGTGCAATGCT	1080
Db	1021	ATTATCAACCCGGAAAGAACCATTCATCTGTAAAGGGCACGTTGAGGCTGTGTGCAATGCT	1080
Qy	1081	GAGATTAAGATTAATGAAGAAGCTGGCGGAGAGCCTTTGAATAATGATAGCTGGCTGTTAAC	1140
Db	1081	GAGATTAAGATTAATGAAGAAGCTGGCGGAGAGCCTTTGAATAATGATAGCTGGCTGTTAAC	1140
Qy	1141	ACCACACTCCGAGATCTCTCCAGCCTGTACCCCATCCACAGTTTGGCCCGCTCCGCAAT	1200
Db	1141	ACCACACTCCGAGATCTCTCCAGCCTGTACCCCATCCACAGTTTGGCCCGCTCCGCAAT	1200
Qy	1201	CATCATCTCTTAATCCAGAGCAGAGAGATTGTGATCTCTTCAATCCCAACCCAGGCTGTGGGG	1260
Db	1201	CATCATCTCTTAATCCAGAGCAGAGAGATTGTGATCTCTTCAATCCCAACCCAGGCTGTGGGG	1260
Qy	1261	GCCATCATCCGGAAAGAAAGGGGGGCACATCAAAACAGCTGGCGAGATTGCGCGAGACTCT	1320
Db	1261	GCCATCATCCGGAAAGAAAGGGGGGCACATCAAAACAGCTGGCGAGATTGCGCGAGACTCT	1320
Qy	1321	ATCAAGATTGGCCCTCGCGAAGAGGCCCAAGCTCCAGCAAAAGATGTCATCATCCCGGG	1380
Db	1321	ATCAAGATTGGCCCTCGCGAAGAGGCCCAAGCTCCAGCAAAAGATGTCATCATCCCGGG	1380
Qy	1381	CCACCGGAAGCCCACTTCAAGGCCCGAGGGACGATCTTTGGGAAACTGAAAGAGGAAAAAC	1440
Db	1381	CCACCGGAAGCCCACTTCAAGGCCCGAGGGACGATCTTTGGGAAACTGAAAGAGGAAAAAC	1440
Qy	1441	TTCTTTAAACCCCAAGAAAGAAAGTGAAGTGAAGGCAATTCAGAGGCCCTCTTCCACA	1500
Db	1441	TTCTTTAAACCCCAAGAAAGAAAGTGAAGTGAAGGCAATTCAGAGGCCCTCTTCCACA	1500
Qy	1501	GCTGGCCGGGTGATTTGGCAAGGTGGCAAGCCGTGAACGAACTGCAGAACTTAACCAAT	1560
Db	1501	GCTGGCCGGGTGATTTGGCAAGGTGGCAAGCCGTGAACGAACTGCAGAACTTAACCAAT	1560
Qy	1561	GCAGAAATCATGTCCTCTGTGACCAAAACGCCAGATGAATAATGAGGAAGTATCTGTACA	1620
Db	1561	GCAGAAATCATGTCCTCTGTGACCAAAACGCCAGATGAATAATGAGGAAGTATCTGTACA	1620
Qy	1621	ATTATCGGGCACTCTTCTTGCTAGCAGACTGACAGCCCAAGATCAGGGAAATTTGTACA	1680
Db	1621	ATTATCGGGCACTCTTCTTGCTAGCAGACTGACAGCCCAAGATCAGGGAAATTTGTACA	1680
Qy	1681	CAGGTAAACAGCAGAGCAGAAATACCTCAGGAGTGTGCTCTCAACGCGCAGCAAGTGA	1740
Db	1681	CAGGTAAACAGCAGAGCAGAAATACCTCAGGAGTGTGCTCTCAACGCGCAGCAAGTGA	1740
Qy	1741	GGCTCCACAGGCAACCGAAGAAACAAACGAGATGATGTAGCCCTTTCGAACCTTGAACAGAA	1800
Db	1741	GGCTCCACAGGCAACCGAAGAAACAAACGAGATGATGTAGCCCTTTCGAACCTTGAACAGAA	1800
Qy	1801	TGAGAACCAACCGCAGCAGCAGATCCGAGAGCAAAACCAAAACCATCTGAGGAATGAGAA	1860
Db	1801	TGAGAACCAACCGCAGCAGCAGATCCGAGAGCAAAACCAAAACCATCTGAGGAATGAGAA	1860
Qy	1861	GTTGTGCGAGCGCGCCAGGGAATCTTCCGAGGCTGTGAGAAACCCAGGGGCGAGAGAGG	1920
Db	1861	GTTGTGCGAGCGCGCCAGGGAATCTTCCGAGGCTGTGAGAAACCCAGGGGCGAGAGAGG	1920
Qy	1921	GCAGGGAAGTCAAGCAGGATTTTGCAGAAACCAACGAGCCGCGCTTCCGCGCCCTCAAGGGC	1980

[illegible]

QY 3061 CCTGACACCTGCTCTGTTGTTCCCTAGAGATTTTGTAAAGCTGATGTTGAGCAT 3120
Db 3061 CCTGACACCTGCTCTGTTGTTCCCTAGAGATTTTGTAAAGCTGATGTTGAGCAT 3120
QY 3121 TTTTATATTTTATATATAAATGAGTTGAGAAAAAATAGATATCAATGCCAGCTTG 3180
Db 3121 TTTTATATTTTATATATAAATGAGTTGAGAAAAAATAGATATCAATGCCAGCTTG 3180
QY 3181 GAGAAAGTGACAGTCCAGTGTGCAACAGCTGTTCTGATGTTCTCCCTAGACCAAGA 3240
Db 3181 GAGAAAGTGACAGTCCAGTGTGCAACAGCTGTTCTGATGTTCTCCCTAGACCAAGA 3240
QY 3241 CCNATATGCGCTCTTTTGACAAACCTTGAATAATGTTATTT 3283
Db 3241 CCNATATGCGCTCTTTTGACAAACCTTGAATAATGTTATTT 3283

RESULT 4
LOCUS AF057352 3667 bp mRNA linear PRI 23-MAY-1999
DEFINITION Homo sapiens hepatocellular carcinoma autoantigen (p62) mRNA,
ACCESSION AF057352
VERSION AF057352
KEYWORDS complete cds.
SOURCE AF057352.1 GI:4883680
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3667)
AUTHORS Zhang, J.Y., Chan, E.K., Peng, X.X. and Tan, E.M.
TITLE A novel cytoplasmic protein with RNA-binding motifs is an
autoantigen in human hepatocellular carcinoma
J. Exp. Med. 189 (7), 1101-1110 (1999)
MEDLINE 99207072
PUBMED 10190901
REFERENCE 2 (bases 1 to 3667)
AUTHORS Zhang, J.Y., Chan, E.K.L., Peng, X.X. and Tan, E.M.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1998) Molecular & Experimental Medicine, The
Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla,
CA 92037, USA

FEATURES

source

1. 3667
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 3667
/gene="p62"
436.. 2106
/gene="p62"
/note="RNA-binding protein"
/codon_start=1
/product="hepatocellular carcinoma autoantigen"
/protein_id="A0J31596.1"
/db_xref="GI:4883681"

gene

CDS

1. 3667
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. 3667
/gene="p62"
436.. 2106
/gene="p62"
/note="RNA-binding protein"
/codon_start=1
/product="hepatocellular carcinoma autoantigen"
/protein_id="A0J31596.1"
/db_xref="GI:4883681"
/translation="MMKKYTGSLSPAVTADDRQLPFGDRKFLPAGCVLTKSGYAVD
YPLDMALIRLETLSGKVEHAGKIMEVDVSVKRSRKIQINIPHLQWELDGL
AQQYGVENEQVETDLETAVNTYATREAKTAMEKLSCHQFENAFKISYLPDEEV
SSPSPPQARQGHSSREQGHAPGTSOAOIDPRLIIVPTQFVAGIIEKGLITKN
ITKOTOSRVDIHRKENSAGAEKPVTTIHAPEGTSACRMLEIMOREADDETKLAETP
IKTLANGIAGRLIKEGRRLKKIEHETGKRTISLSLODSTYNPERTTVKTVKAC
ASAETIMKLRARAFENDMLAVNTHSGYRSLSLPHQGFPPHHSYPEQETVNLTP
TQAVGATIGKGAHITQOLAFAGASIKIAEGLPDSERVVITGPEAOFKQGIIF
GKLENFVPEKKEVRLBAHRIVSSTAGVIGKGTVEILQNLTSAEIVTRDQTP
DENEEIVIRLIIGHFASQTAQRKIRIVQVKQOQKYPQGVASQSK"

ORIGIN

Query Match 99.1%; Score 3253.8; DB 9; Length 3667;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3277; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 3 CAGCGAGAGAGAGGAGGAGCGCGGGGTACCGGGCCGGGGAGCCCGGGGCTCTCGGGGAA 62
Db 369 CAGCGAGAGAGGAGGAGCGCGGGGTACCGGGCCGGGGAGCCCGGGGCTCTCGGGGAA 428
QY 63 GAGACGATGATGAACAAGCTTTTACATCGGGAACTTGAGCCCGCGCTCAACGCGAGGA 122
Db 429 GAGACGATGATGAACAAGCTTTTACATCGGGAACTTGAGCCCGCGCTCAACGCGAGGA 488
QY 123 CTTCCGAGACTCTTTTGGGGAACAGAACTGCCCCCTGGCGGGGACAGTCTGCTGAAGTC 182
Db 489 CTTCCGAGACTCTTTTGGGGAACAGAACTGCCCCCTGGCGGGGACAGTCTGCTGAAGTC 548
QY 183 CGGCTAGCGCTTCTGAGACTACCCCGAACCAAACTGGGCGCATTCGCGCATGAGACCT 242
Db 549 CGGCTAGCGCTTCTGAGACTACCCCGAACCAAACTGGGCGCATTCGCGCATGAGACCT 608
QY 243 CTGCGGTAAAGTGAATTCATGAGGAAATCATGGAAGTGAATTCATGAGTCTGTAATAA 302
Db 609 CTGCGGTAAAGTGAATTCATGAGGAAATCATGGAAGTGAATTCATGAGTCTGTAATAA 668
QY 303 GCTAAGAGAGAGGAAATTCATGAGTTCGAAACATCCCTCTGACCTGAGTGGAGTGT 362
Db 669 GCTAAGAGAGAGGAAATTCATGAGTTCGAAACATCCCTCTGACCTGAGTGGAGTGT 728
QY 363 GGATGSACTTTTGGCTCAATATGGGAGAGTGGAGAAATGTTGGAACAAGTCAACAGACAC 422
Db 729 GGATGSACTTTTGGCTCAATATGGGAGAGTGGAGAAATGTTGGAACAAGTCAACAGACAC 788
QY 423 AGAAACCGCGCTTTCACGTCATATGCAACAGAGAGAGAGCAAAATAGCCATGGA 482
Db 789 AGAAACCGCGCTTTCACGTCATATGCAACAGAGAGAGAGCAAAATAGCCATGGA 848
QY 483 GAAGCTAAGCGGGCATGATTTGGAACATCTCCTTCAAGATTTCTATCATCCGGATGA 542
Db 849 GAAGCTAAGCGGGCATGATTTGGAACATCTCCTTCAAGATTTCTATCATCCGGATGA 908
QY 543 AGAGGTAGGTCCTCCCTGCGCCCTCGACGAGCCACGCTGGGGACCACTTCCCGGGA 602
Db 909 AGAGGTAGGTCCTCCCTGCGCCCTCGACGAGCCACGCTGGGGACCACTTCCCGGGA 968
QY 603 GCAAGGCGACGCGCTTGGGGGCACTTCTGAGGCGAGACAGATTGATTTCCCGTGGGAT 662
Db 969 GCAAGGCGACGCGCTTGGGGGCACTTCTGAGGCGAGACAGATTGATTTCCCGTGGGAT 1028
QY 663 CTGCTCCCAACCCAGTTTGTGGTGCATCTGGAAGAGGCGCTTGAACATTAAGAA 722
Db 1029 CTGCTCCCAACCCAGTTTGTGGTGCATCTGGAAGAGGCGCTTGAACATTAAGAA 1088
QY 723 CATCACTAAGCAGACCCAGTCCCGGGTATGATTCATGAAAAAGAACTTGAAGCTGC 782
Db 1089 CATCACTAAGCAGACCCAGTCCCGGGTATGATTCATGAAAAAGAACTTGAAGCTGC 1148
QY 783 AGAGAGGCGCTGACCATCCATCCACCCAGAGGGGACTTCTGAACATGCGCATGAT 842
Db 1149 AGAGAGGCGCTGACCATCCATCCACCCAGAGGGGACTTCTGAACATGCGCATGAT 1208
QY 843 TCTTGAATATATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 902
Db 1209 TCTTGAATATATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1268
QY 903 AATCTTGGGACAAAGTGGCTTGTGGAGAGCTGATTTGGAAGAAAGAGGAGGAGGAGGAGGAG 962
Db 1269 AATCTTGGGACAAAGTGGCTTGTGGAGAGCTGATTTGGAAGAAAGAGGAGGAGGAGGAGGAG 1328
QY 963 GAAATATGAACATGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1022
Db 1329 GAAATATGAACATGAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1388
QY 1023 ATACAAACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1082
Db 1389 ATACAAACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1448
QY 1083 GATAGAGATTAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1142

1449 GATGAGATTATGAGAGAGCTGCGTAGAGCCCTTTGAAAATGATATGCTGGCTTTAAAC 1508
1143 CCACTCCGGATATCTTCACAGCTGTACCCCATCAGAGTTGGCCGTTCCGCATCA 1202
1509 CCACTCCGGATATCTTCACAGCTGTACCCCATCAGAGTTGGCCGTTCCGCATCA 1568
1203 TCACCTTATCCAGAGAGAGATTGTGATCTTCATCCCAACCCAGGCGTGGGGGC 1262
1569 TCACCTTATCCAGAGAGAGATTGTGATCTTCATCCCAACCCAGGCGTGGGGGC 1628
1263 CATCATCGGAAAGAGAGGCGGCAACATCAAAACAGCTGGAGATTTGCGCGAGCCCTAT 1322
1629 CATCATCGGAAAGAGAGGCGGCAACATCAAAACAGCTGGAGATTTGCGCGAGCCCTAT 1688
1323 CAAGATTGCCCCCTGGGAAAGCCCAAGCTGACGAAAGAGATGTATATCAACGGGCG 1382
1689 CAAGATTGCCCCCTGGGAAAGCCCAAGCTGACGAAAGAGATGTATATCAACGGGCG 1748
1383 ACCGGAAGCCCATGTTCAAGGCCCAAGGACGGATCTTTGGGAAACTGAAAGAGAAACTT 1442
1749 ACCGGAAGCCCATGTTCAAGGCCCAAGGACGGATCTTTGGGAAACTGAAAGAGAAACTT 1808
1443 CTTTAAACCCCAAGAAAGAGTGAAGCTGGAGCGCATATCAGAGTGCCCTCTTCCACAGC 1502
1809 CTTTAAACCCCAAGAAAGAGTGAAGCTGGAGCGCATATCAGAGTGCCCTCTTCCACAGC 1868
1503 TGGCCGCGGTGATTGGCAAAAGTGGCAAGCCGTGAACGAATGAGAAATGAGAAATGAGT 1562
1869 TGGCCGCGGTGATTGGCAAAAGTGGCAAGCCGTGAACGAATGAGAAATGAGAAATGAGT 1928
1563 AAGAGTCATGTCGCTCGTGACCAACCGCAGATGAAATGAGAAATGAGAAATGAGT 1622
1929 AAGAGTCATGTCGCTCGTGACCAACCGCAGATGAAATGAGAAATGAGAAATGAGT 1988
1623 TATCGGCGCACTCTTGTAGCCAGACTGACAGCGCAAGATCAGGGAATTTGTACACA 1682
1989 TATCGGCGCACTCTTGTAGCCAGACTGACAGCGCAAGATCAGGGAATTTGTACACA 2048
1683 GGTGAAGCAGCAGAGCAGAAATACCTTACAGGAGTGGCTTCAAGCGCAGCAAGTGAAG 1742
2049 GGTGAAGCAGCAGAGCAGAAATACCTTACAGGAGTGGCTTCAAGCGCAGCAAGTGAAG 2108
1743 CTCGCCAAGGCGCCAGCAAAACAAAGGATGATGATCCCTTCAACACTGACAGATG 1802
2109 CTCGCCAAGGCGCCAGCAAAACAAAGGATGATGATGATCCCTTCAACACTGACAGATG 2168
1803 AGACCAAGCAGCAGCAGAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAAAT 1862
2169 AGACCAAGCAGCAGCAGCAGAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAAAT 2228
1863 CTCGGAAGCGCGCAGAGACTCTGCGAGGCGCTTGAGAACCCCAAGGGCGCAGAGAGGGC 1922
2229 CTCGGAAGCGCGCAGAGACTCTGCGAGGCGCTTGAGAACCCCAAGGGCGCAGAGAGGGC 2288
1923 GGGGAAAGTTCAGCAGGTTTGGCAGAACCAACGAGGCCCGCTCCGCGCCCAAGGGCTT 1982
2289 GGGGAAAGTTCAGCAGGTTTGGCAGAACCAACGAGGCCCGCTCCGCGCCCAAGGGCTT 2348
1983 CTCGAGGCTTCAGCCATCCACTTCAACCATCACTCGGATCTCTCTGAACCTCCACAGAG 2042
2349 CTCGAGGCTTCAGCCATCCACTTCAACCATCACTCGGATCTCTCTGAACCTCCACAGAG 2408
2043 CTATCCCTTTTATGTTGACATCAATAGTGAAGCTGTTCAAAAGCCAAAGAAATGACAC 2102
2409 CTATCCCTTTTATGTTGACATCAATAGTGAAGCTGTTCAAAAGCCAAAGAAATGACAC 2468
2103 CCTTTTCTGAGCAATGCTCTCTGTACATGTTGTACATATTTAGAAAGGAGATGTT 2162
2469 CCTTTTCTGAGCAATGCTCTCTGTACATGTTGTACATATTTAGAAAGGAGATGTT 2528
2163 AAGATATGTCCTGTGGTTTACACAGGTCCTGCAAGCGGTATATATTTTGAAGAAATA 2222

2529 AAGATATGTCCTGTGGTTTACACAGGTCCTGCAAGCGGTATATATTTTGAAGAAATA 2588
2223 TATATCAAAATACCAACTTAACTCCAAATTTTATCAATTAATTAATTTTCTTTT 2282
2589 TATATCAAAATACCAACTTAACTCCAAATTTTATCAATTAATTAATTTTCTTTT 2648
2283 AAAGAGAAAGCAGGCTTTTCTAGAATTTAAAGAAATTAAGTCTTTGGAGGCTCAGCGTG 2342
2649 AAAGAGAAAGCAGGCTTTTCTAGAATTTAAAGAAATTAAGTCTTTGGAGGCTCAGCGTG 2708
2343 TAGAGAGAGCTTTGAGGCGACCCGCAAAAATTCACCAAGGAAATCTCGTGGAG 2402
2709 TAGAGAGAGCTTTGAGGCGACCCGCAAAAATTCACCAAGGAAATCTCGTGGAG 2768
2403 GACACTCAGGCAAGTTCTGATCACTGTGTATGTCAACAGAGGATTAACGCTCTCTTG 2462
2769 GACACTCAGGCAAGTTCTGATCACTGTGTATGTCAACAGAGGATTAACGCTCTCTTG 2828
2463 AAGAGAAAGCTGTCTGATCTGTATGCTGTATGCTATACACCATTTCTTTGCTT 2522
2829 AAGAGAAAGCTGTCTGATCTGTATGCTGTATGCTATACACCATTTCTTTGCTT 2888
2523 CACAGTTTAACTGTTTTTGAATCTGCTATATATATCTGTCTCTCTGTTA 2582
2889 CACAGTTTAACTGTTTTTGAATCTGCTATATATATCTGTCTCTCTGTTA 2948
2583 TCTCTCCCTCCCTCCCTCCCTCTTCTCCATCTCCATCTTTTGAATTTCTGTATCC 2642
2949 TCTCTCCCTCCCTCCCTCCCTCTTCTCCATCTCCATCTTTTGAATTTCTGTATCC 3008
2643 CTCATCTCATCTCGGTATCTTACGACCCCGCCCGCCAGGCAAGAGTCTGAT 2702
3009 CTCATCTCATCTCGGTATCTTACGACCCCGCCCGCCAGGCAAGAGTCTGAT 3067
2703 ATCAGTCAACAAGAAAGAAAGCAAGCAACAAACAGCTCACTTACACTTGA 2762
3069 ATCAGTCAACAAGAAAGAAAGCAAGCAACAAACAGCTCACTTACACTTGA 3127
2763 TTAATCAAAAGAAAGAAAGTCAATGTTACTTGTCTGAGGTTTGGAGAGAAACAG 2822
3128 TTAATCAAAAGAAAGAAAGTCAATGTTACTTGTCTGAGGTTTGGAGAGAAACAG 3187
2823 AACCCAAACCAACCAATCAACCAACCAAGAAATTTCCAAATTAAGAAATGAT 2882
3188 AACCCAAACCAACCAATCAACCAACCAAGAAATTTCCAAATTAAGAAATGAT 3247
2883 TTTGCTTTTGTGATTTGATGATTAAGCCATCAATATCAGCAAAATGATCTTCTT 2942
3248 TTTGCTTTTGTGATTTGATGATTAAGCCATCAATATCAGCAAAATGATCTTCTT 3307
2943 T-AAAAAAGAAAGTGAAGAAAGTGAATTTTCAAGGTTTGGCCAGAGCGTTA 3001
3308 TAAAAAAGAAAGTGAAGAAAGTGAATTTTCAAGGTTTGGCCAGAGCGTTA 3367
3002 AATTCAAGATTTTATACAGAAAAACACAGAGAAAGTCAACTCAGGTGTTTATC 3061
3368 AATTCAAGATTTTATACAGAAAAACACAGAGAAAGTCAACTCAGGTGTTTATC 3427
3062 CTCAGACCTGCTCTGTGTTTCCCTTAGAGATTTGTAAGCTGATAGTGAAGAT 3121
3428 CTCAGACCTGCTCTGTGTTTCCCTTAGAGATTTGTAAGCTGATAGTGAAGAT 3487
3122 TTTTATTTTTTATTAAGAAAGTGAAGAAATTAAGATATCACTGCGAGCTGG 3181
3488 TTTTATTTTTTATTAAGAAAGTGAAGAAATTAAGATATCACTGCGAGCTGG 3547
3182 AGAAGGTGACAGTCCAGAGTGAACAGCTGTTCTGAATGCTTCGCGTACGCAAGAC 3241
3548 AGAAGGTGACAGTCCAGAGTGAACAGCTGTTCTGAATGCTTCGCGTACGCAAGAC 3607
3242 CNAATAGGCTCTGTTTGAACAACTTGAAGATTTATTT 3283
3608 CTATATGCTCTCTTTTGAACAACTTGAAGATTTATTT 3649

Db	1741	ATGCTCAGAAATTATCGGGCACTTCTTTGCTAGCCACACTGACACAGCGCAAAATCAGGGAA	1800
Qy	1672	ATTGTATCAACAGGTGAAGACGACGAGAGACAGAAATACCTTCAGGGAGTGGCTTCACAGCGC	1731
Db	1801	ATTGTATCAACAGGTGAAGACGACGAGAGAGAAATACCTTCAGGGAGTGGCTTCACAGCGCC	1860
Qy	1732	AGCAAGTGAAGCTCCCAAGGCACGAGCACGAGCAAAACAAGGATGTATAGGCTTCCAAAC	1791
Db	1861	AGCAAGTGAAGCTCCCAAGGCACGAGCACGAGCAAAACAAGGATGTATAGGCTTCCAAAC	1920
Qy	1792	CTGACAGATGAGACCAACGACGACGACGACGATCGGGAGCAAAACCAAGACCATCTGAG	1851
Db	1921	CTGACAGATGAGACCAACGACGACGACGACGATCGGGAGCAAAACCAAGACCATCTGAG	1980
Qy	1852	GAATGAGAAGTCTGCGAGAGCGGCGGACGAGGACTCTGCCGAGGCCCTGAGAACCCACGGGC	1911
Db	1981	GAATGAGAAGTCTGCGAGAGCGGCGGACGAGGACTCTGCCGAGGCCCTGAGAACCCACGGGC	2040
Qy	1912	CGAGAGAGGGGCGGGGAAGTCAAGCAGATTGTGCAGAACCAACGAGGCCGCGCTCCGGCC	1971
Db	2041	CGAGAGAGGGGCGGGGAAGTCAAGCAGATTGTGCAGAACCAACGAGGCCGCGCTCCGGCC	2100
Qy	1972	CCCCAGGGCTTCTGACGGCTTCAGCCATCCACTCACTCCAGACTCTGCGATCTCTCTGAA	2031
Db	2101	CCCCAGGGCTTCTGACGGCTTCAGCCATCCACTCACTCCAGACTCTCTCTGAA	2160
Qy	2032	CTCCACAGACGCTATCCCTTTTGTAGTAACTAACATAGTGAAGTGTTCAAAAGCCAGC	2091
Db	2161	CTCCACAGACGCTATCCCTTTTGTAGTAACTAACATAGTGAAGTGTTCAAAAGCCAGC	2220
Qy	2092	AAATATGACACCCCTTTTCTGTGGCAATGCTCTGTATACATGTGTATCATATTGAA	2151
Db	2221	AAATATGACACCCCTTTTCTGTGGCAATGCTCTGTATACATGTGTATCATATTGAA	2280
Qy	2152	GGGAAGATGTTAAGTATGTGGCCGTGGGGTTACAGAGGTGCTGACGCGTATATAT	2211
Db	2281	GGGAAGATGTTAAGTATGTGGCCGTGGGGTTACAGAGGTGCTGACGCGTATATAT	2340
Qy	2212	TTTAAATAATATATATCAAAATPACTCAACTAGTCCAAATTTTATCAATTATTAATTTT	2271
Db	2341	TTTAAATAATATATATCAAAATPACTCAACTAGTCCAAATTTTATCAATTATTAATTTT	2400
Qy	2272	TTTTTCTTTTAAAGAAAGCAGGCTTTTCTAATTTTAAAGAAATAAAGCTTTGGGAG	2331
Db	2401	TTTTTCTTTTAAAGAAAGCAGGCTTTTCTAATTTTAAAGAAATAAAGCTTTGGGAG	2460
Qy	2332	GCTCACGCTGTAGAGAGAGGCTTTGAGGCCACCGGCAAAATTCAACCCAGAGGGAAT	2391
Db	2461	GCTCACGCTGTAGAGAGAGGCTTTGAGGCCACCGGCAAAATTCAACCCAGAGGGAAT	2520
Qy	2392	CTCGTCGGAAGACACTCAGCGCAGTTGTGATCACTGTATGTATCAAGAGGGATA	2451
Db	2521	CTCGTCGGAAGACACTCAGCGCAGTTGTGATCACTGTATGTATCAAGAGGGATA	2580
Qy	2452	CCGTCCTCTTGAAGAGAACTCTGTCACTCTCATAGCTGTCTAAGTCAACCCAT	2511
Db	2581	CCGTCCTCTTGAAGAGAACTCTGTCACTCTCATAGCTGTCTAAGTCAACCCAT	2640
Qy	2512	TCTCTTGTCTTACAGGTTTTTAACTGTTTTTTCATACATGCTATATATTTCTCTGCT	2571
Db	2641	TCTCTTGTCTTACAGGTTTTTAACTGTTTTTTCATACATGCTATATATTTCTCTGCT	2700
Qy	2572	CTCTCGTTATCTCTCCCTCCCTCCCTCCCTCTCTTCATCTCCATTTCTTTTGA	2631
Db	2701	CTCTCGTTATCTCTCCCTCCCTCCCTCCCTCTCTTCATCTCCATTTCTTTTGA	2760
Qy	2632	TTTTCCTCATCCCTCATCTCAATCCCGTATCTAGCACCCGCCGCCGCCGAGCAAGCA	2691
Db	2761	TTTTCCTCATCCCTCATCTCAATCCCGTATCTAGCACCCGCCGCCGCCGAGCAAGCA	2820
Qy	2692	GTCCTCTGATATCATATACCAAAAGGAACAAAAGGAAACACAGAAACAGCTCAA	2751
Db	2821	GTCCTCTGATATCATATACCAAAAGGAACAAAAGGAAACACAGAAACAGCTCAA	2880

Oy		2752	CTTACACTGGTTGATACATAAAGAAACAAGAAGTCAAATGTACTCTTGAGCGTTTGGAG	2811
Dd		2881	CTTACACTGGTTGATACATAAAGAACCAAGACTCAATGTACTCTTGAGCGTTTGGAG	2940
Oy		2812	AGGAAGAAACAGAAACCACCAACCAACCATCAACCAACCAAGAAAAATTCCACATG	2871
Dd		2941	AGGAAGAAACAGAAACCACCAACCAACCAATCAACCAACCAAGAAAAATTCCACATG	3000
Oy		2872	AAAAGATGTATTTTGTCTTTTGGCATTTTGTGTATTAAGCATTCAATATTCAGCAAAATG	2931
Dd		3001	AAAAGATGTATTTTGTCTTTTGCATTTTGGTGTATTAAGCATTCAATATTCAGCAAAATG	3060
Oy		2932	ATTCCCTTCTTTAAAAAAAATATGTGAGAGAAAGTAGAATTTTACCAAGTTGTGGCC	2991
Dd		3061	ATTCCCTTCTTTAAAAAAAATATGTGAGAGAAAGTAGAATTTTACCAAGTTGTGGCC	3120
Oy		2992	CAGGCGCTTAATTCACAGATTTTTTTTACGAGAAAAACACAGAGAAAGCTACCTCAG	3051
Dd		3121	CAGGCGCTTAATTCACAGATTTTTTTTAAAGAAAAACACAGAGAAAGCTACCTCAG	3180
Oy		3052	GTTGTTTTTACCTCAGGACCTTGCTCTGTGTCTTCCCTTAGATTTTGTAAAGCTGATAG	3111
Dd		3181	GTTGTTTTTACCTCAGGACCTTGCTCTGTGTCTTCCCTTAGAGATTTTGTAAAGCTGATAG	3240
Oy		3112	TTCGAGCATTTTTTTTATTTTTTAAATAAATAGATGTGAAAAAAAATAAGATATCACT	3171
Dd		3241	TTCGAGCATTTTTTTTATTTTTTAAATAAATAGATGTGAAAAAAAATAAGATATCACT	3300
Oy		3172	GCCAGCCTGAGAAAGTGTCAGTCACAGTGTGCACACAGCTGTTCTGAAATGTCTCCGCT	3231
Dd		3301	GCCAGCCTGAGAAAGTGTCAGTCACAGTGTGCACACAGCTGTTCTGAAATGTCTCCGCT	3360
Oy		3232	AGCCAAGAACCMATATGAGCGCTCTCTTTTGGACAAACCTTTGAAATGTTATTT	3283
Dd		3361	AGCCAAGAACCMATATGAGCGCTCTCTTTTGGACAAACCTTTGAAATGTTATTT	3412
RESULT 6 AR343074				
LOCUS	AR343074	3412 bp	DNA	linear PAT 17-AUG-2003
DEFINITION	Sequence 6 from patent US 6576756.			
ACCESSION	AR343074			
VERSION	AR343074.1	GI:3738476		
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 3412) Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K. and Old,L.J.			
TITLE	Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof			
JOURNAL	Patent: US 6576756-A 6 10-JUN-2003;			
FEATURES	Location/Qualifiers			
source	1..3412			
ORIGIN	/organism="unknown" /mol_type="genomic DNA"			
Query Match	95.7%; Score 3143; DB 6; Length 3412;			
Best Local Similarity	96.2%; Pred. No. 0;			
Matches 32B3; Conservative	0; Mismatches 0; Indels 129; Gaps 1;			
Oy		1	GGCACGGAGAGAGCGGAGAGCGCGGGTACCGGGCCGAGGAGGCCGCGGAGCTCTCGGGG	60
Dd		1	GGCACGGAGAGAGCGGAGAGCGCGGGTACCGGGCCGAGGAGGCCGCGGAGCTCTCGGGG	60
Oy		61	AAAGAGCGGATGATGAAACAAGCTTTTACATCGGGAACCTGAGCCCCCGCGCTCAACGCGCAC	120
Dd		61	AAAGAGCGGATGATGAAACAAGCTTTTACATCGGGAACCTGAGCCCCCGCGCTCAACGCGCAC	120
Oy		121	GAACCTCCGACACTCTTTTGGGACAAGGAAGCTGCCCTTGGCGGACAAGGTCTTGCTGAAG	180

Db	121	GA	CTCCGGAGAGCTCTTTGGGGACAGAAAGCTGCTCCCTGGCGGGACAGGTCCTGCTGAAG	180
Qy	181	TC	GGCTAGCGCTTTCGTGACTACCCCGACAGAACTGGGCAATCCGCGCATCGAGACC	240
Db	181	TC	GGCTAGCGCTTTCGTGACTACCCCGACAGAACTGGGCAATCCGCGCATCGAGACC	240
Qy	241	CT	CTCGGGGTAAAGTGGAAATTCGATGGGAAATATATGAGAAAGTTGATTAATCTAGTCTCTAAA	300
Db	241	CT	CTCGGGGTAAAGTGGAAATTCGATGGGAAATATATGAGAAAGTTGATTAATCTAGTCTCTAAA	300
Qy	301	AA	GCTTAAGGACAGAAAAATTCAGATTCGAAACATCCCTCCCTCACTGGACGTGGAGAGG	360
Db	301	AA	GCTTAAGGACAGAAAAATTCAGATTCGAAACATCCCTCCCTCACTGGACGTGGAGAGG	360
Qy	361	TT	GGATGACTTTTGGCTCAATATGGGACAGTGGAAATGTGGAAACAGTCAACACAGAC	420
Db	361	TT	GGATGACTTTTGGCTCAATATGGGACAGTGGAAATGTGGAAACAGTCAACACAGAC	420
Qy	421	AC	AGAAACCGCGGTGTCAAGTCACATATTCGAAACAAGAAAGAAACAAAAATAGCATG	480
Db	421	AC	AGAAACCGCGGTGTCAAGTCACATATTCGAAACAAGAAAGAAACAAAAATAGCATG	480
Qy	481	GAG	AAGCTAAGCGGCGCATCAGTTTGAAGACTACTCCTTCAAGATTTCCATCATCCCGAT	540
Db	481	GAG	AAGCTAAGCGGCGCATCAGTTTGAAGACTACTCCTTCAAGATTTCCATCATCCCGAT	540
Qy	541	GA	AGAGGTGAGCTCCCTCTTGCCTCCCTCAGCGAGCCAGCGTGGGGACCATCTTCCCGG	600
Db	541	GA	AGAGGTGAGCTCCCTCTTGCCTCCCTCAGCGAGCCAGCGTGGGGACCATCTTCCCGG	600
Qy	601	GAG	CAAGGCGACAGCCCTGGGGGGACCTTCTAGGGCGACACAAATTTGATTTCCGCTGGGG	660
Db	601	GAG	CAAGGCGACAGCCCTGGGGGGACCTTCTAGGGCGACACAAATTTGATTTCCGCTGGGG	660
Qy	661	AT	CTCGATCCCAACCCAGTTTGTGTGGCAATCATCGGAAAGGAGGGCTTGAACAATAAG	720
Db	661	AT	CTCGATCCCAACCCAGTTTGTGTGGCAATCATCGGAAAGGAGGGCTTGAACAATAAG	720
Qy	721	AA	CACTAAGCAGACCCAGTCCCGGGTATTCATATCCATAGAAAAGAAACTCTGAGACT	780
Db	721	AA	CACTAAGCAGACCCAGTCCCGGGTATTCATATCCATAGAAAAGAAACTCTGAGACT	780
Qy	781	GC	AGGAAGCCGTGACCATCCAGCCGAGGGGACCTTGAAGCATGCGCGATG	840
Db	781	GC	AGGAAGCCGTGACCATCCAGCCGAGGGGACCTTGAAGCATGCGCGATG	840
Qy	841	AT	CTTGAATCATGACAGAAAGGCGCATGAGACCAACTAGCCGAGAGATTCCTGTG	900
Db	841	AT	CTTGAATCATGACAGAAAGGCGCATGAGACCAACTAGCCGAGAGATTCCTGTG	900
Qy	901	AA	AATCTTGGCACCAATGGCTTGTGTAAGCTGATTTGAAAAAGAGGCGAGAAATTTG	960
Db	901	AA	AATCTTGGCACCAATGGCTTGTGTAAGCTGATTTGAAAAAGAGGCGAGAAATTTG	960
Qy	961	AA	AAAAATTGAAACATGAGGCTTGTGTAAGCTGATTTGAAAAAGAGGCGAGAAATTTG	1020
Db	961	AA	AAAAATTGAAACATGAGGCTTGTGTAAGCTGATTTGAAAAAGAGGCGAGAAATTTG	1020
Qy	1021	AT	ATATCAACCCGGAAGAACCATCATCTGTGTAAGGGCACAGTTGAGGCTGTGCGCATG	1080
Db	1021	AT	ATATCAACCCGGAAGAACCATCATCTGTGTAAGGGCACAGTTGAGGCTGTGCGCATG	1080
Qy	1081	GAG	ATATGAGATTAATGAAGAAAGCTGCGTGAAGGCTTTGAAAATGATATGCTGGCTGTAAC	1140
Db	1081	GAG	ATATGAGATTAATGAAGAAAGCTGCGTGAAGGCTTTGAAAATGATATGCTGGCTGTAAC	1140
Qy	1141	CA	ACAGGCAATCTGATCCAGGGTGAACCTCAGCGCATCTTGGCATCTTTTCAACAGGA	1200
Db	1141	CA	ACAGGCAATCTGATCCAGGGTGAACCTCAGCGCATCTTGGCATCTTTTCAACAGGA	1200
Qy	1141	---	---	1140

Db	1201	CTGTCCGTGCTATCTTCCACAGCAGAGGCCCGCGGAGGCTCCCGCGCTGCCCTTACAC	1260
QY	1141	-----ACCCACCTCCGGAACTCTTCCAGCCGTGACCCCATCACAGATTGGCCCG	1197
Db	1261	CCCTTCACTACCACTCCGGAACTTCTCCAGCCGTGACCCCATCACAGATTGGCCCG	1320
QY	1192	TTCCCGCATCATCACTCTTATCCAGAGAGAGATTGGAATCTTTATCCCAACCCG	1251
Db	1321	TTCCCGCATCATCACTCTTATCCAGAGAGAGATTGGAATCTTTATCCCAACCCG	1380
QY	1252	GCTGTGGGCGCATCATGCGGAGAGAGGGGCGCACATCAACAGCTGGAGATTGCGC	1311
Db	1381	GGTGTGGGCGCATCATGCGGAGAGAGGGGCGCACATCAACAGCTGGAGATTGCGC	1440
QY	1312	GGAGGCTCTATCAAGATTGGCCCTGGGAAAGGCCCGACGTTCAGGAAAGGATGGTCA	1377
Db	1441	GGAGGCTCTATCAAGATTGGCCCTGGGAAAGGCCCGACGTTCAGGAAAGGATGGTCA	1500
QY	1372	ATCACCGGCGCCACCGGAAGCCAGTTCAGAGCCAGGACCGATCTTTGGGAACTGAA	1431
Db	1501	ATCACCGGCGCCACCGGAAGCCAGTTCAGAGCCAGGACCGATCTTTGGGAACTGAA	1560
QY	1432	GAGGAAAACTTCTTTAACTCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGATGCC	1491
Db	1561	GAGGAAAACTTCTTTAACTCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGATGCC	1620
QY	1492	TCTTCAACAGCTGGCGCGGTGATTTGGCAAGGTGGCAAGCCGTGACGAATCCGACAC	1551
Db	1621	TCTTCAACAGCTGGCGCGGTGATTTGGCAAGGTGGCAAGCCGTGACGAATCCGACAC	1680
QY	1552	TTAACAGATGAGAGAGTCAATCGTCCCTGTGACCAACCGCAGATGGAATGAGGAAGTG	1611
Db	1681	TTAACAGATGAGAGAGTCAATCGTCCCTGTGACCAACCGCAGATGGAATGAGGAAGTG	1740
QY	1612	ATCGTCAGATTTATCGGCGACTTCTTTGCTAGCCAGACTGTCACAGCCGAAATCAGGAA	1671
Db	1741	ATCGTCAGATTTATCGGCGACTTCTTTGCTAGCCAGACTGTCACAGCCGAAATCAGGAA	1800
QY	1672	ATTGTAACAAGGTGGAGAGACAGACAGAGCAAAATACCTTCAGGGAGTCCGCTCACAGCC	1731
Db	1801	ATTGTAACAAGGTGGAGAGACAGACAGAGCAAAATACCTTCAGGGAGTCCGCTCACAGCC	1860
QY	1732	AGCAAGTGAAGCTCCACAGGCAACAGCAAAAACAAGCATGAATGTAGCCCTTCAACAC	1791
Db	1861	AGCAAGTGAAGCTCCACAGGCAACAGCAAAAACAAGCATGAATGTAGCCCTTCAACAC	1920
QY	1792	CTGACAGATGAGACCAACCGCAGCCAGCCAGATCGGAGCAAAACCAAGACATCTGAG	1851
Db	1921	CTGACAGATGAGACCAACCGCAGCCAGCCAGATCGGAGCAAAACCAAGACATCTGAG	1980
QY	1852	GAAAGAGAGGTGCGGAGCGGAGCGGCGCAGGGAATCTGTGCGAAGCCCTTGAAACCCAGGGGC	1911
Db	1981	GAAAGAGAGGTGCGGAGCGGAGCGGCGCAGGGAATCTGTGCGAAGCCCTTGAAACCCAGGGGC	2040
QY	1912	CGAGGAGGGGCGGAGGAGTCAAGCAGGTTTGCACAGAACCAACGAGGCCCGCTCCGCGC	1971
Db	2041	CGAGGAGGGGCGGAGGAGTCAAGCAGGTTTGCACAGAACCAACGAGGCCCGCTCCGCGC	2100
QY	1972	CCCCAGGCTTCTGACAGGCTTTCAGCCATTCACATTCACCTCGATCTCTCTGAA	2031
Db	2101	CCCCAGGCTTCTGACAGGCTTTCAGCCATTCACATTCACCTCGATCTCTCTGAA	2160
QY	2032	CTCCACAGAGCGTATCCCTTTTAGTTGAATTAACATATGAGGAAGTGTTCAAAGCCAGC	2091
Db	2161	CTCCACAGAGCGTATCCCTTTTAGTTGAATTAACATATGAGGAAGTGTTCAAAGCCAGC	2220
QY	2092	AAATGACACCCCTTTTCTGTGGCAAAATGCTCTGTGATCATGCTGTGATCATTTAGAA	2151
Db	2221	AAATGACACCCCTTTTCTGTGGCAAAATGCTCTGTGATCATGCTGTGATCATTTAGAA	2280
QY	2152	GGGAGAGATGTTAAGATATGTGCGCTGTGGGTTACACAGGCTGCTCAGCCGATATATAT	2211
Db	2281	GGGAGAGATGTTAAGATATGTGCGCTGTGGGTTACACAGGCTGCTCAGCCGATATATAT	2340

QY 2212 TTGAAATATATATCAATCACTCACTCAATTTTATCAATTTATTTT 2271
 Db 2241 TTGAAATATATATCAATCACTCACTCAATTTTATCAATTTATTTT 2400
 QY 2272 TTTTCTTTTAAAGAGAGAGAGCTTTCTAGACTTTTAAAGATTTTGGAG 2331
 Db 2401 TTTTCTTTTAAAGAGAGAGAGCTTTCTAGACTTTTAAAGATTTTGGAG 2460
 QY 2332 GTCTCAGGCTGAGAGAGAGCTTTAGAGCCACCCGCAAAATTTACCCAGAGGAAT 2391
 Db 2461 GTCTCAGGCTGAGAGAGAGCTTTAGAGCCACCCGCAAAATTTACCCAGAGGAAT 2520
 QY 2392 CTCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2451
 Db 2521 CTCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
 QY 2452 CCGTCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2511
 Db 2581 CCGTCTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
 QY 2512 TCTCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2571
 Db 2641 TCTCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
 QY 2572 CTCTCTGTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2631
 Db 2701 CTCTCTGTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
 QY 2632 TTTTCTTCT 2691
 Db 2761 TTTTCTTCT 2820
 QY 2692 GTGCTCTGAGATCAATCACTCACTCACTCACTCACTCACTCACTCA 2751
 Db 2821 GTGCTCTGAGATCAATCACTCACTCACTCACTCACTCACTCACTCA 2880
 QY 2752 CTTCACCTGTTATCTCTCAAAAGAGAGAGAGAGAGAGAGAGAGAG 2811
 Db 2881 CTTCACCTGTTATCTCTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
 QY 2812 AGGAAAG 2871
 Db 2941 AGGAAAG 3000
 QY 2872 AAGAAATGATTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCT 2931
 Db 3001 AAGAAATGATTTTGTCTTTTGTCTTTTGTCTTTTGTCTTTTGTCT 3060
 QY 2932 ATTCTCTTCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2991
 Db 3061 ATTCTCTTCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
 QY 2992 CAGGAGGTTAAATTCACAGATTTTAAAGAGAGAGAGAGAGAGAGAGAG 3051
 Db 3121 CAGGAGGTTAAATTCACAGATTTTAAAGAGAGAGAGAGAGAGAGAGAG 3180
 QY 3052 GTGTTTTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3111
 Db 3181 GTGTTTTTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
 QY 3112 TTGAGAGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 3171
 Db 3241 TTGAGAGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 3300
 QY 3172 GCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3231
 Db 3301 GCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
 QY 3232 AGCCAAAG 3283
 Db 3361 AGCCAAAG 3412

RESULT 7
 BD209925 3412 bp DNA linear PAT 17-JUL-2003
 LOCUS BD209925
 DEFINITION Isolated nucleic acid molecules encoding cancer-associated antigens, these antigens and method of using the same.
 ACCESSION BD209925
 VERSION BD209925.1 GI:33019695
 KEYWORDS JP 2002512049-A/4.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3412)
 AUTHORS Chen, Y.T., Gure, A., Tsang, S., Stockert, E., Jager, E., Knuth, A. and Old, L.J.
 TITLE Isolated nucleic acid molecules encoding cancer-associated antigens, these antigens and method of using the same
 JOURNAL Patent: JP 2002512049-A 4 23-APR-2002;
 LUDWIG INSTITUTE FOR CANCER RESEARCH
 COMMENT OS Homo sapiens (human)
 PN JP 2002512049-A/4
 PD 23-APR-2002
 PF 16-MAR-1999 JP 2000545030
 PR 17-APR-1998 US 09/061709
 PI YAO TSENG CHEN, ALI GURE, SOLAM TSANG, ELISABETH STOCKERT, ELKE JAGER,
 PI ALEXANDER KNUTH, LOYD J OLD
 PC C12N15/09, A61K35/12, A61K39/00, A61P35/00, C07K16/32, PC C12N1/15,
 PC C12N1/19, C12N1/21, C12N5/10, C12P21/08, C12O1/68, G01N33/53, G01N33/574,
 C12N15/00, C12N5/00
 CC n-not determined(3372)
 FH Key Location/Qualifiers
 FT source 1..3412
 FT Location/Qualifiers
 FT /organism='Homo sapiens (human)'.
 FT /mol_type='genomic DNA'
 FT /db_xref='taxon:9606'
 ORIGIN
 Query Match 95.7%; Score 3143; DB 6; Length 3412;
 Best Local Similarity 96.2%; Pred. No. 0;
 Matches 3283; Conservative 0; Mismatches 0; Indels 129; Gaps 1;
 QY 1 GGCAGCGAG 60
 Db 1 GGCAGCGAG 60
 QY 61 AAGAGAGAGATGATGAAACAGCTTTACATCGGGGAACTGAGCCCGGCTGACCGCGAG 120
 Db 61 AAGAGAGAGATGATGAAACAGCTTTACATCGGGGAACTGAGCCCGGCTGACCGCGAG 120
 QY 121 GACCTCCGGAGAGCTCTTTGGGGAGCAGAGAGCTCCCTGCGGAGACAGTCTGTTGAAG 180
 Db 121 GACCTCCGGAGAGCTCTTTGGGGAGCAGAGAGCTCCCTGCGGAGACAGTCTGTTGAAG 180
 QY 181 TCCGGCTACGCTTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 Db 181 TCCGGCTACGCTTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 QY 241 CTCTCGGTTAAAGTGAATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 Db 241 CTCTCGGTTAAAGTGAATTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
 QY 301 AAGCTAAG 360
 Db 301 AAGCTAAG 360

QY 361 TTGATGACCTTTTGCTCATATGAGACAGTGAAGATGTGAAACAAGTCAACACAGAC 420
DB 361 TTGATGACCTTTTGCTCATATGAGACAGTGAAGATGTGAAACAAGTCAACACAGAC 420
QY 421 AAGAAACCGCGCTGTGTCACTGCATATGCAACAGAAAGAAACAAATAAGCCATG 480
DB 421 AAGAAACCGCGCTGTGTCACTGCATATGCAACAGAAAGAAACAAATAAGCCATG 480
QY 481 GAGAAGCTAAGCGGCGATCAGTTTGAAACTAATCTCTTCAAAATTTTCCATATCCCGAT 540
DB 481 GAGAAGCTAAGCGGCGATCAGTTTGAAACTAATCTCTTCAAAATTTTCCATATCCCGAT 540
QY 541 GAAAGAGTGAAGCTCCCTTTCGCGCCCTCAGCGAGCCAGCGTGGGAGACCACTTCGCGG 600
DB 541 GAAAGAGTGAAGCTCCCTTTCGCGCCCTCAGCGAGCCAGCGTGGGAGACCACTTCGCGG 600
QY 601 GAGCAGGCGCAGCGCCCTTGGGGGCACTTCTCAAGGCGAAGATGATTTTCCGCTGCGG 660
DB 601 GAGCAGGCGCAGCGCCCTTGGGGGCACTTCTCAAGGCGAAGATGATTTTCCGCTGCGG 660
QY 661 ATCTGCTGCCCAACCCAGTTTGTGTGTCATCATCGGAAAGAGGGCTTGACATTAAG 720
DB 661 ATCTGCTGCCCAACCCAGTTTGTGTGTCATCATCGGAAAGAGGGCTTGACATTAAG 720
QY 721 AACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAAAGAACTTGGAGCT 780
DB 721 AACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAAAGAACTTGGAGCT 780
QY 781 GCAGAGAAAGCTGTGACCATCATATGCCACCCGAGGGGCACTTGAAGATGCCGATG 840
DB 781 GCAGAGAAAGCTGTGACCATCATATGCCACCCGAGGGGCACTTGAAGATGCCGATG 840
QY 841 ATTCTGAAATCATGCGAAGAGCGAGATGAGACCAAACTAGCGGAAGAAATTCCTGTG 900
DB 841 ATTCTGAAATCATGCGAAGAGCGAGATGAGACCAAACTAGCGGAAGAAATTCCTGTG 900
QY 901 AAAATCTTGGCACAACAATGCTTGTGGAAGACTGATTTGAAAAAGAGCAGAAATTTG 960
DB 901 AAAATCTTGGCACAACAATGCTTGTGGAAGACTGATTTGAAAAAGAGCAGAAATTTG 960
QY 961 AAGAAATTTGAACATGAAACAGGACCAAGATTAACAATCTGATCTTTGCAAGATTTGAGC 1020
DB 961 AAGAAATTTGAACATGAAACAGGACCAAGATTAACAATCTGATCTTTGCAAGATTTGAGC 1020
QY 1021 ATATACAAACCCGAAAGAACCATCATCTGTGAAGGGCACAATTGAGGCTGTGCGAGTGT 1080
DB 1021 ATATACAAACCCGAAAGAACCATCATCTGTGAAGGGCACAATTGAGGCTGTGCGAGTGT 1080
QY 1081 GAGATGAGATTAAGAAAGAGCTGCGAGGCTTTGAAAAATGATATGCTGTGTTAAC 1140
DB 1081 GAGATGAGATTAAGAAAGAGCTGCGAGGCTTTGAAAAATGATATGCTGTGTTAAC 1140
QY 1141 ----- 1140
DB 1141 CAAACAGCAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200
QY 1141 ----- 1140
DB 1201 CTGTCCGTCTATCTTCCACGACGAGGCCCGCGAGAGCTCCCCCGCTGCCCTACAC 1260
QY 1141 -----ACCACTCCGANAATTTCTCCAGCTGTACCCCAATCACAAGTTTGGCCG 1191
DB 1261 CCGTTCACTAACCACTCCGATTACTTCTCCAGCTGTACCCCAATCACAAGTTTGGCCG 1320
QY 1192 TTCCCGCATCATCATCTTTATCCAGAGCAGAGATTTGAAATCTCTTCATCCCAACCCAG 1251
DB 1321 TTCCCGCATCATCATCTTTATCCAGAGCAGAGATTTGAAATCTCTTCATCCCAACCCAG 1380
QY 1252 GCTGTGGGGCGCATCATCCGGAAGAAAGGGGGCACAATCAAAACAGCTGGGGAATTTGCC 1311
DB 1381 GCTGTGGGGCGCATCATCCGGAAGAAAGGGGGCACAATCAAAACAGCTGGGGAATTTGCC 1440
QY 1312 GAGGCTCTATCAAGATTTGCCCTTGGCGAGAGGCCCAAGAGCTCAGCGAAAGATGTCATG 1371

DB 1441 GAGGCTCTATCAAAATTTGCCCTTGGCGAGAGGCCCAAGAGCTCAGCGAAAGATGTCATC 1500
QY 1372 ATCAACGGGCGCACCGGAAGCCAGTTCAAGGCCCAAGGACGAGATCTTTGGGAACTGAA 1431
DB 1501 ATCAACGGGCGCACCGGAAGCCAGTTCAAGGCCCAAGGACGAGATCTTTGGGAACTGAA 1560
QY 1432 GAGGAAATCTCTTTAAACCCCAAGAAAGATGAAGCTGGAAGCCATATCAGAGTCCC 1491
DB 1561 GAGGAAATCTCTTTAAACCCCAAGAAAGATGAAGCTGGAAGCCATATCAGAGTCCC 1620
QY 1492 TCTTCAAGCTGGCGGGGTGATTTGGCAAGGTGGCAAGACCTGTAACGAATCGCAAGAC 1551
DB 1621 TCTTCAAGCTGGCGGGGTGATTTGGCAAGGTGGCAAGACCTGTAACGAATCGCAAGAC 1680
QY 1552 TTAACAGATGAGAAATCATGTCGCTGCTGTCGCAAGCCAGATGAAATAGGAAGTG 1611
DB 1681 TTAACAGATGAGAAATCATGTCGCTGCTGTCGCAAGCCAGATGAAATAGGAAGTG 1740
QY 1612 ATCTGAGAAATTAATCGGCACTTCTTTGCTAGCCAGACTGCAAGCGCAAGATCAGGAA 1671
DB 1741 ATCTGAGAAATTAATCGGCACTTCTTTGCTAGCCAGACTGCAAGCGCAAGATCAGGAA 1800
QY 1672 ATGTGACAAAGGTGAAGACGAGAGGAATTAATCCTCAGGAGTGCCTCACAAGGC 1731
DB 1801 ATGTGACAAAGGTGAAGACGAGAGGAATTAATCCTCAGGAGTGCCTCACAAGGC 1860
QY 1732 AGCAAGTGAAGCTCCACAGGACCAAGCAAAACAAGATGAATGATCCCTTCCAAACAG 1791
DB 1861 AGCAAGTGAAGCTCCACAGGACCAAGCAAAACAAGATGAATGATCCCTTCCAAACAG 1920
QY 1792 CTGACAGATGAGACCAACCGCAGCCAGCCAGATCGGAGCAAAACAAAGACATCTGAG 1851
DB 1921 CTGACAGATGAGACCAACCGCAGCCAGCCAGATCGGAGCAAAACAAAGACATCTGAG 1980
QY 1852 GAATGAGAAAGTCTGCGAGAGGCGGCAAGGACTCTGCGAGGCTGTGAGAACCCCAAGGGC 1911
DB 1981 GAATGAGAAAGTCTGCGAGAGGCGGCAAGGACTCTGCGAGGCTGTGAGAACCCCAAGGGC 2040
QY 1912 CGAAGAGGGGCGGGGAAGTCAAGCAGGTTTGCAGAAACCAAGAGCCGCTCCGCGC 1971
DB 2041 CGAAGAGGGGCGGGGAAGTCAAGCAGGTTTGCAGAAACCAAGAGCCGCTCCGCGC 2100
QY 1972 CCCCAGGGCTTCTGCAAGCTTCAAGCCATCACTTCAACATCACTCGAGTCTCTGAA 2031
DB 2101 CCCCAGGGCTTCTGCAAGCTTCAAGCCATCACTTCAACATCACTCGAGTCTCTGAA 2160
QY 2032 CTCCACAGAGGCTATCCCTTTAGTTGAACATAATGAGTGAAGCTGTCAAGCCAAAGC 2091
DB 2161 CTCCACAGAGGCTATCCCTTTAGTTGAACATAATGAGTGAAGCTGTCAAGCCAAAGC 2220
QY 2092 AAAATGACACCCCTTTTCTGTGGCAAAATCGTCTGTACATGTGTATCAATTAAGA 2151
DB 2221 AAAATGACACCCCTTTTCTGTGGCAAAATCGTCTGTACATGTGTATTAAGA 2280
QY 2152 GGGAGATGTTAAGATATGTGCCCTGTGGTTACAGAGGTCCTGACGGCGATATAT 2211
DB 2281 GGGAGATGTTAAGATATGTGCCCTGTGGTTACAGAGGTCCTGACGGCGATATAT 2340
QY 2212 TTTAGAAATATATATCAATTAACCTCAACCTCAATTTTAAATTAATTAATTTT 2271
DB 2341 TTTAGAAATATATATCAATTAACCTCAACCTCAATTTTAAATTAATTAATTTT 2400
QY 2272 TTTTCTTTTAAAGAAAGAGGCTTTCTAGACTTTTAAAGAAATTAAGCTTTGGAG 2331
DB 2401 TTTTCTTTTAAAGAAAGAGGCTTTCTAGACTTTTAAAGAAATTAAGCTTTGGAG 2460
QY 2332 GTCTCAAGGTGAGAGAGGCTTTGAGGCAACCCGACAAATTCACCCAGAGGAAT 2391
DB 2461 GTCTCAAGGTGAGAGAGGCTTTGAGGCAACCCGACAAATTCACCCAGAGGAAT 2520
QY 2392 CTGCTGGAAGAGCACTCAAGGAGTTCTGATCACTGTGTATGTCACAGAAAGGATA 2451

```

Db      2521 CTGCTGGAGAGACACTCAGCGAGTTCTGTGATCACCCTGTATGTCAACGAGGATTA 2580
Qy      2452 CCGTCTCCTTGAAGAGAACTCTGTGCACTCTCTCAATGCTGTCTATGCTCAATACCCATT 2511
Db      2581 CCGTCTCCTTGAAGAGAACTCTGTGCACTCTCTCAATGCTGTCTATGCTCAATACCCATT 2640
Qy      2512 TCTCTGCTTCAACGAGTTTAAATGATTTTTCATACGACCTATATATCTCTGCT 2571
Db      2641 TCTCTGCTTCAACGAGTTTAAATGATTTTTCATACGACCTATATATCTCTGCT 2700
Qy      2572 CTCTGTGTTATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2631
Db      2701 CTCTGTGTTATCTCTCCCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCT 2760
Qy      2632 TTTCTCTATCCCTCATCTCAATCCCGTATCTAGGACACCCCTCCCTCCCTCCCTCCCT 2691
Db      2761 TTTCTCTATCCCTCATCTCAATCCCGTATCTAGGACACCCCTCCCTCCCTCCCTCCCT 2820
Qy      2692 GTGCTGTGATTCATCATCATCAACAAAGAGCAAAAGCAACACCAACAGCCTCA 2751
Db      2821 GTGCTGTGATTCATCATCATCAACAAAGAGCAAAAGCAACACCAACAGCCTCA 2880
Qy      2752 CTTCACCTTGTGTTACTCAAAAGACAGATGTAATGTAATGTAATGTAATGTAATGTA 2811
Db      2881 CTTCACCTTGTGTTACTCAAAAGACAGATGTAATGTAATGTAATGTAATGTAATGTA 2940
Qy      2812 AGGAAAAAGAGAACCCCAACCAACCAATCAACCAACCAACCAACCAACCAACCAATG 2871
Db      2941 AGGAAAAAGAGAACCCCAACCAACCAATCAACCAACCAACCAACCAACCAACCAATG 3000
Qy      2872 AAGAAATGATTTTGTCTTTTGTGATTTTGTGATTAACCATCAATTCATGACAAATG 2931
Db      3001 AAGAAATGATTTTGTCTTTTGTGATTTTGTGATTAACCATCAATTCATGACAAATG 3060
Qy      2932 ATTCCTTTCTTTAAAAAAAATGTGAGAGAAAGTGAATTTACCAAGTTGTGGCC 2991
Db      3061 ATTCCTTTCTTTAAAAAAAATGTGAGAGAAAGTGAATTTACCAAGTTGTGGCC 3120
Qy      2992 CAGGCGGTAAATTCACATTTTAAACGAAAAACACAGAGAAAGTCACTCAG 3051
Db      3121 CAGGCGGTAAATTCACATTTTAAACGAAAAACACAGAGAAAGTCACTCAG 3180
Qy      3052 GTGTTTAACTGACGACCTGTCTGTGTTTCCCTTGAAGATTTTGAAGCTGATG 3111
Db      3181 GTGTTTAACTGACGACCTGTCTGTGTTTCCCTTGAAGATTTTGAAGCTGATG 3240
Qy      3112 TTGAGACATTTTAAATTTTAAATGAATGAATGAATGAATGAATGAATGAATGAAT 3171
Db      3241 TTGAGACATTTTAAATTTTAAATGAATGAATGAATGAATGAATGAATGAATGAAT 3300
Qy      3172 GCCAGCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3231
Db      3301 GCCAGCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3360
Qy      3232 AGCCAGAAACCAATATGAGCTCTTTTGAACAAACCTTGAATTTTATTT 3283
Db      3361 AGCCAGAAACCAATATGAGCTCTTTTGAACAAACCTTGAATTTTATTT 3412

```

```

RESULT 8
BC021290 3633 bp mRNA linear PRI 04-OCT-2003
LOCUS     Homo sapiens IGF-II mRNA-binding protein 2, mRNA (cDNA clone
DEFINITION
MGC:29539 IMAGE:5090334), complete cds.
ACCESSION BC021290
VERSION    BC021290.2 GI:33878041
KEYWORDS   MGC.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3633)
AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,

```

REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED

2 (bases 1 to 3633)
Strausberg,R.
Direct Submission
Submitted (14-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK COMMENT

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:18204200.
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

FEATURES

source

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/INL at: <http://image.llnl.gov>
Series: IRAL Plate: 39 Row: 0 Column: 21
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 27552765.
Location/Qualifiers

gene

1..3633
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:29539 IMAGE:5090334"
/tissue_type="Pancreas, epithelioid carcinoma"
/clone_lib="NIH MGC 42"
/lab_host="DH10B-R"
/note="Vector: POTB7"

CDS

1..3633
/gene="IMP-2"
/note="synonym: VICKZ2"
/db_xref="LocustID:10644"
68..1864
/codon_start=1
/product="IGF-II mRNA-binding protein 2"
/protein_id="AAH21290.1"

[illegible]

Db 1741 CAGAAATTAACGGGCACTTTCTGACGACACTGACAGCGCAAGATCAGGAAATTTGT 1800
 QY 1677 ACAACAGGTGAAGACGACGAGGACAGAAATACCTTCAGGAGTCTGCTCAACGCGCAGAA 1736
 Db 1801 ACAACAGGTGAAGACGACGAGGACAGAAATACCTTCAGGAGTCTGCTCAACGCGCAGAA 1860
 QY 1737 GTGAGGCTTCCCAACGAGGACGAGGAAACAAAGGATGAAATGATAGCCTTCCAAACCTGAG 1796
 Db 1861 GTGAGGCTTCCCAACGAGGACGAGGAAACAAAGGATGAAATGATAGCCTTCCAAACCTGAG 1920
 QY 1797 AGAATGAGACCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1856
 Db 1921 AGAATGAGACCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1980
 QY 1857 AGAATGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1916
 Db 1981 AGAATGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2040
 QY 1917 AGGAGCGGGGAAAGTCAAGGACGAGTTGGCAGAAACGACGACGACGACGACGACGACGACG 1976
 Db 2041 AGGAGCGGGGAAAGTCAAGGACGAGTTGGCAGAAACGACGACGACGACGACGACGACGACG 2100
 QY 1977 GGGCTTCTGACGAGCTTCAAGCCTCACTTCAACATTCACCTCGATCTCTGAACTCCG 2036
 Db 2101 GGGCTTCTGACGAGCTTCAAGCCTCACTTCAACATTCACCTCGATCTCTGAACTCCG 2160
 QY 2037 ACAGCGCTATCCCTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 2096
 Db 2161 ACAGCGCTATCCCTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 2220
 QY 2097 GAGACGCCCTTTTCTGAGCAATGCTCTGTAACATGCTGTAACATGCTGTAACATGCTGTAAC 2156
 Db 2221 GAGACGCCCTTTTCTGAGCAATGCTCTGTAACATGCTGTAACATGCTGTAACATGCTGTAAC 2280
 QY 2157 GATGTTAAGATATGTCGCTGAGGTTTACACAGGTCGCTGACAGCGGTAATATTTTAC 2216
 Db 2281 GATGTTAAGATATGTCGCTGAGGTTTACACAGGTCGCTGACAGCGGTAATATTTTAC 2340
 QY 2217 AATATATATATCAATTAATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2276
 Db 2341 AATATATATATCAATTAATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2400
 QY 2277 CTTTAAAGAGAAACGAGCTTTTCAAGCTTAAAGATTAAGCTTTTGGAGGCTTC 2236
 Db 2401 CTTTAAAGAGAAACGAGCTTTTCAAGCTTAAAGATTAAGCTTTTGGAGGCTTC 2460
 QY 2337 ACAGGTGAGAGAGAGCTTTGAGGCAACCGCAAAATTCACCCAGAGGAAATCTCGT 2396
 Db 2461 ACAGGTGAGAGAGAGCTTTGAGGCAACCGCAAAATTCACCCAGAGGAAATCTCGT 2520
 QY 2397 CGGAGAGACCTCAACGAGCTTTCGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTC 2456
 Db 2521 CGGAGAGACCTCAACGAGCTTTCGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTC 2580
 QY 2457 TCCCTTGAAGAGAACTCTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2516
 Db 2581 TCCCTTGAAGAGAACTCTGCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
 QY 2517 TTGCTTCAAGGTTTAACTGTTTAACTGTTTAACTGTTTAACTGTTTAACTGTTTAACTGTTTAACT 2576
 Db 2641 TTGCTTCAAGGTTTAACTGTTTAACTGTTTAACTGTTTAACTGTTTAACTGTTTAACTGTTTAACT 2700
 QY 2577 TGTATTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2636
 Db 2701 TGTATTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760
 QY 2637 TCAATCCCTCATCTCAATCCCTCATCTCAATCCCTCATCTCAATCCCTCATCTCAATCCCTCATCT 2696
 Db 2761 TCAATCCCTCATCTCAATCCCTCATCTCAATCCCTCATCTCAATCCCTCATCTCAATCCCTCATCT 2819
 QY 2697 CTGAGTATCATCATCAACAAAGAGACAAAGGCAAAACACAAACACAGCTCTCAACTTAC 2756
 Db 2820 CTGAGTATCATCATCAACAAAGAGACAAAGGCAAAACACAAACACAGCTCTCAACTTAC 2879

QY 2757 ACTTGTTACTCAAAAGACAGAGTCATGTAATTGTCCTAGCGTTTGGAGAGAA 2816
 Db 2880 ACTTGTTACTCAAAAGACAGAGTCATGTAATTGTCCTAGCGTTTGGAGAGAA 2939
 QY 2817 AACAGAAACCAACCAACCAACCAATCAATCAACCAACCAACCAACCAATCAATCAATCAATCAAT 2876
 Db 2940 AACAGAAACCAACCAACCAACCAATCAATCAACCAACCAACCAACCAATCAATCAATCAATCAAT 2999
 QY 2877 ATGATATTTGCTTTTTCATTTTGTGTATTAAGCATCAATATTCACCAAAATGATTC 2936
 Db 3000 ATGATATTTGCTTTTTCATTTTGTGTATTAAGCATCAATATTCACCAAAATGATTC 3059
 QY 2937 TTTCTTTT-AAAAAAAAAATGAGAGAGTAGAAATTTTACCAAGTTTGGCCACG 2995
 Db 3060 TTTCTTTTAAAAAAAAAATGAGAGAGTAGAAATTTTACCAAGTTTGGCCACG 3119
 QY 2996 GCGTTAAATTCACAGATTTTAAACGAGAAACACACAGAGAGAGTCACTCAGGT 3055
 Db 3120 GCGTTAAATTCACAGATTTTAAACGAGAAACACACAGAGAGAGTCACTCAGGT 3179
 QY 3056 TTTTACCTCAGACCTTGCTCTGTTGTTCCCTTACAGATTTTGTAAAGCTATAGTTG 3115
 Db 3180 TTTTACCTCAGACCTTGCTCTGTTGTTCCCTTACAGATTTTGTAAAGCTATAGTTG 3239
 QY 3116 AGCATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 3175
 Db 3240 AGCATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 3299
 QY 3176 GCGTGAAGAGGTGACAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 3235
 Db 3300 GCGTGAAGAGGTGACAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTC 3359
 QY 3236 AAGAACCAATATGAGCTTCTTTTGGACAAACCTTGAATTTTATTT 3283
 Db 3360 AAGAACCAATATGAGCTTCTTTTGGACAAACCTTGAATTTTATTT 3407

RESULT 9
 AC020629 142971 bp DNA linear PRI 07-MNR-2002
 LOCUS Homo sapiens 12q BAC Rpl1-7E16 (Hoswell Park Cancer Institute
 DEFINITION Human BAC library) complete sequence.
 ACCESSION AC020629
 VERSION AC020629.6 GI:765675
 KEYWORDS HMG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 142971)
 REFERENCE
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Muzny,D.M., Adams,C., Bailey,M., Barbarta,J., Blankenburg,K.,
 Bodoc,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
 Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
 David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
 Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferriguto,D.,
 Forcum-Tansey,J., Frantz,P., Ganesu,R., Gorrell,J.H., Gorrell,L.L.,
 Guevara,W., Harris,K., Hernandez,J., Hodgson,A., Hogue,M.,
 Hollway,C., Hosak,H., Issar,A., Jackson,J.E., Jackson,L., Jia,Y.,
 Jones,M., Kelly,S., Kneitz,S., Kondajewski,N., Kong,Y., Kovar,C.,
 Lau,S., Leal,B., Lee,E., Li,Z., Lichtarge,O., Liu,J., Liu,M.,
 Logan,O., Lozano,R.J., Lu,J., Lucier,R., Marondel,I., Martin,R.,
 Martinez,C., McLeod,M.P., Mel,G., Merscher,S., Nelson,A.,
 Montgomery,K.T., Morgan,M., Morris,S., Naah,S., Nelson,S.,
 Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B., Paxton,S.,
 Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D., Rives,M.,
 Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Shim,C., Simon,M.,
 Sparks,A., Stamps,A., Sugang,R., Tabot,P., Taylor,T., Vasquez,L.,
 Vanson,R., Vo,O., Wahbah,M., Watlington,S., Weinstein,G.,
 Weinstein,I.R., Williamson,A., Worley,K., Wren,J., Wrensford,G.,
 Gibbs,A.M., Yang,R., Yu,W., Zhou,X., Kuchelapatti,R., Nelson,D. and
 Direct Submission

TITLE

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 142971)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 142971)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 4 (bases 1 to 142971)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (28-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 5 (bases 1 to 142971)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Apr 27, 2000 this sequence version replaced gi:7025656.
COMMENT INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-14) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 142971
Phrap values in estimate: 140680
Average error rate (BCM-Phrap estimate): 0.000118703
Fraction of Phrap values less than 40 : 0.0433608
Number of consensus changing edits: 31
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
11171 gtctctcga (n) caaaaaaa gtctctcga (a) caaaaaaa
12018 tactatctc (n) ctatctatc tactatctc (t) ctatctatc
16604 tccactctaa (n) agtccattt tccactctaa (g) agtccattt
18385 tggtaaccac (n) taccatttta tggtaaccac (t) taccatttta
21879 aaaaaaaaaa (n) nttaaacag aaaaaaaaaa (a) nttaaacag
21880 aaaaaaaaaa (n) nttaaacag aaaaaaaaaa (a) nttaaacag
21881 aaaaaaaaaa (n) nttaaacag aaaaaaaaaa (a) nttaaacag
21906 cctaggtccc (n) cttaaatgag cctaggtccc (a) cttaaatgag
38215 accagctcgt (n) caaatgagc accagctcgt (c) caaatgagc
38915 ttacatctag (n) tcttctcaaa ttacatctag (t) tcttctcaaa
55675 acttctctc (n) aaaaaaaaaa acttctctc (a) aaaaaaaaaa
57658 gaccagctgt (n) ccaaatgagc gaccagctgt (a) ccaaatgagc
74371 gtctcggccc (n) ccgcccggc gtctcggccc (c) ccgcccggc
88021 taaatgagg (n) agaccatgtg taaatgagg (c) agaccatgtg
88248 ggaatgagc (n) agatcacac ggaatgagc (a) agatcacac
94121 ttgtccctga (n) gcttcctgc ttgtccctga (t) gcttcctgc
95228 gtcacatgtc (n) cmmttgttc gtcacatgtc (t) cmmttgttc
95230 ccattgtgnc (n) nttgttcaa ccattgtgnc (t) nttgttcaa
95231 catgtgnc (n) ntgttcaac catgtgnc (c) ntgttcaac
95232 atgtgncm (n) ttgttcaact atgtgncm (t) ttgttcaact
104931 tttttttag (n) accgaagggt tttttttag (t) accgaagggt
118168 gcccaagctc (n) ccaaatgagc gcccaagctc (c) ccaaatgagc
120974 catctcaaaa (n) anaaagagc catctcaaaa (a) anaaagagc
120976 tctcaaaaaa (n) anaaagagc tctcaaaaaa (a) anaaagagc
120978 tcaaaaana (n) anaaagagc tcaaaaana (a) anaaagagc
129580 aaaaanaana (n) gaaactcttc aaaaanaana (a) gaaactcttc
139505 ccagcatagg (n) gggggagc ccagcatagg (t) gggggagc
139508 gccatgggg (n) gggcaccgt gccatgggg (t) gggcaccgt
139508 ataccatca (n) aatatgtag ataccatca (g) aatatgtag
140104 accgtgccca (n) cagaatgaga accgtgccca (g) cagaatgaga
140149 tggaccatct (n) gataccocaa tggaccatct (g) gataccocaa

----- Distribution of Quality < 40 Bases -----

#	bases	5	10	15	20	25	30	35	40
1000	900	*	*	*	*	*	*	*	*
800	800	*	*	*	*	*	*	*	*
700	700	*	*	*	*	*	*	*	*
600	600	*	*	*	*	*	*	*	*
500	500	*	*	*	*	*	*	*	*
400	400	*	*	*	*	*	*	*	*
300	300	*	*	*	*	*	*	*	*
200	200	*	*	*	*	*	*	*	*
100	100	*	*	*	*	*	*	*	*
0	0	*	*	*	*	*	*	*	*

Phrap Value Range

Version: 1.01 gxt0.
FEATURES
source Location/Qualifiers
1..142971
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-76B16"
147..236
/rpt_family="A-rich"
repeat_region
258..378
/rpt_family="AluB"
repeat_region
complement(537..840)
/rpt_family="AluX"
repeat_region
complement(1722..2137)
/rpt_family="Tiger2a"
repeat_region
complement(3161..3584)

QY	1801	TAGAACCAACGACGACGACGATCGGGAGCAACCAACCAATCTGAGAAATGAGAA	1860
Db	134732	TGAGACCAACCAACGACGACGATCAAGAGCAACCAACCAATCTGAGAAATGAGAA	134732
QY	1861	GTCGCGGAGCGCGCCAGGGACTCTGCAGGCGCCTTAGAAACCCGAGGGCCGAGGAGGG	1920
Db	134792	GTCGCGGAGCGCGCCAGGGACTCTGTAAGGCGCCTTAGAAACCCGAGGGCCGAGGAGGG	134855
QY	1921	GCGGGGAAGTCAAGCCAGGTTTGGCAGAACCAACGAGCCCGCGCTCCGCCCCAGGGC	1980
Db	134852	GTCGGGGAAGGTAAACCAAGGTTTGGCAGAAACCAACGGGCGCCCTCTCTGTCGCCAGGGC	134911
QY	1981	TTGTGAGGGTTTAGGCATCCACATTCACATCCACTGGATCTCTCTGAATCTCCACGA	2040
Db	134912	TTGTGAGGGTTTAGGCATCCACTTCACATCCACTGGATCTCTCTGAATCTCCATGTA	134972
QY	2041	CGCTATCCCTTTAGTTGAACTAACTAAGGTGAAC-GTGTCAAGCCAGCAAAATGCA	2039
Db	134972	CGCTATCCCTTTAGTTGAACTAAATAGTAAACATTTGTCAGAGCCAGCAAAATGCA	135033
QY	2100	CACCCCTTTTCTGTGCAAAATCGTCTCTGTACATGTGTGACATATTAGAAAGGAAGAT	2159
Db	135032	CGCCCTTTT-TGTGCAAAATGCTCTCTACACGTGTATACATATTAGAAAGGAAGAT	135091
QY	2160	GTTAAGATATGTGGCGCTGTGGGTTTACAAGGGGCGCTGCAGCGGTATATTATTTAGAAA	2219
Db	135091	GTTAAGATATGTGGCTGTGGGTTTACAAGGGGCGCTGCAGCGGTATATTATTTAGAAA	135156
QY	2220	TAAATATCAAAATTAATCAACTACTCCAAATTTTATCAATTAATTTTCTTCTT	2279
Db	135151	TAAATATCAAAAT-----AATCACTCCAAATTTTATCAATTAATTA--TTTCTTCTT	135202
QY	2280	TTTAAAGAGAAAGCAGGCTTTTCTAGACTTTAAAGATAAAGTCTTGGAGGCTCACG	2339
Db	135203	TTTAAAGAGAGAGCAGGCTTTTCTAGACTTTAAAGTCT--TTCTTAAAGGCTCACG	135255
QY	2340	GTTGAGAGAGGCTTTGAGGCCACCCGACAAATTTACCCAGAGGAAATCTCGTCG	2399
Db	135260	GTTGAGAGAGGCTTTGAGGCCACCCACAAATTTACCCAGAGGAAATTTGTCGG	135319
QY	2400	AAGGACACTACGAGGAGTTCTGATCACTGTGATGTAAAGAGAGGATACCGTCTCC	2459
Db	135320	AAGGACACTACAGAGATTCTGATACCTATGATGTAAAGAGAGGATCTGCTCC	135379
QY	2460	TTGAAGAGAAACCTGTCTACTCTCATGACCTGTCTAGCTATACCCGATTTCTCTTGS	2519
Db	135380	TTGAAGAGGAAACCTGTCTACTCTCATGCTGCTGTAGGTATACGCCATTTCTCTTGS	135439
QY	2520	CTTACAGGTTTAACTGAGTTTGTGCACTGCTATATTAATTCTCTGTCTCTCTGT	2579
Db	135440	CTTACAGGTTTAACTGAGTTTGTGCACTGCTATATTAATTCTCTGTCTCTCTC--	135496
QY	2580	TTATCTCTCCCTCCCTCCCTCCCTCTTCTCCATCTCCATCTCTTGAATTCCTCA	2639
Db	135497	-TATCTCTCCCTCCCTCCCTCCCTCTTCTCTATCTCCATCTCTTGAATTCCTCA	135555
QY	2640	TCCCTCATCTCAATCCGATCTACGAC-----CCGCCGCCCCAGGCAAGA	2691
Db	135586	TCCCTCATCTCAATCCCATCTATCTAGACCCGCCGACCCGCCCTGCGAGGCAAGA	135615
QY	2692	GTCGCTGAGTATCATCATCAACAAAAGGAACAAAAGCAACACAAACGACTCAA	2751
Db	135616	GTCGCTGAGTATCATCATCAACAAAAGGAACAAAAGCAACACAAACGACTCAA	135675
QY	2752	CTTACACTGTGTTACTCAAAAGAAACAGAGTCAGATGTCCTGTACGTTTGGAG	2811
Db	135676	CTTACACTGTGTTACTCAAAAGAACAGAGTCAGATGTCCTGTCTACATTTGGAA	135735
QY	2812	AGGAAAACAGAACCCACCAACCAACCAATCAACCAACAAAGAAAAATTCACAATG	2871
Db	135736	AGGAAAACAGAACCTATCAACCAACCAACCAATCAACCAACAAAGAAAAATTCACAATG	135795
QY	2872	AAAGATATATTTGCTTTTGGACTTTTGGCTATTAAGCATCAATATTACGCAAAATG	2931

[illegible]

IMPORTANT: This sequence is not the entire insert of clone RP11-325P15. It may be shorter because we sequence overlapping sections only once, except for a short overlap. The true left end of clone RP11-325P15 is at 1 in this sequence. The true left end of clone RP11-337C18 is at 89085 in this sequence. The true right end of clone RP4-704D21 is at 41034 in this sequence.

FEATURES

Source

Location/Qualifiers
1..91084
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-325P15"
/clone_11b="RPCT-11.2"

ORIGIN

Query Match 84.7%; Score 2781.4; DB 9; Length 91084;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 3089; Conservative 0; Mismatches 167; Indels 48; Gaps 16;

```
QY 1 GGCACGCGAGAGGCGCGCGGTAACCGCGCGCGGAGCCCGCGCTTCGCGG 60
DB 9381 GGCACGCGAGAGGCGCGCGGTAACCGCGCGCGGAGCCCGCGCTTCGCGG 9440
QY 61 AAGAACCGGATGATGAACAGCTTTATCATCGGGAACTGAGCCCGCGCTCAACGCGAC 120
DB 9441 AAGAACCGGATGATGAACAGCTTTATCATCGGGAACTGAGCCCGCGCTCAACGCGAC 9500
QY 121 GACCTCCGCGACGCTCTTGGGAGCAGAAAGCTCCCTGCGGAGCAGGTCCTGCTGAAG 180
DB 9501 GACCTCTGGCAGCTCTTGGGAGCAGAAAGCTCCCGCCACGAGCAGGTCCTGCTGAAG 9560
QY 181 TCCGCTACGCTTGTGCTGCTACCTCCGACGAGACTGCGCCATCCGCGCCATCGAGACC 240
DB 9561 TCCGCTACGCTTGTGCTGCTACCTCCGACGAGACTGCGCCATCCGCGCCATCGAGACC 9620
QY 241 CTCTCGGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 9621 CTCTCGGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9680
QY 301 AAGCTAAGAGCAGGAAATTCAGATTCGAAACATCCCTCTCACTGACGTGGAGGTG 360
DB 9681 AAGCTAAGAGCAGGAAATTCAGATTCGAAACATCCCTCTCACTGACGTGGAGGTG 9740
QY 361 TTGATGATGATTTGGCTCAATATGGAACAGTGGAAATGTGG--AACTGTCACAA 417
DB 9741 TTGATGATGATTTGGCTCAATATGGAACAGTGGAAATGTGGAAATGTGGAAATGTGG 9800
QY 418 GACACGAAACCGCGCTGTCACGTCATGCAATGCAACAAAGAGAAATGAGCC 477
DB 9801 GACACGAAACCGCGCTGTCACGTCATGCAATGCAACAAAGAGAAATGAGCC 9855
QY 478 ATGAGAAGCTAAGCGGCGCATCAGTTGAGAACTACTCTTCAAGATTTCTTCAATCCG 537
DB 9856 ATGAGAAGCTAAGCGGCGCATCAGTTGAGAACTACTCTTCAAGATTTCTTCAATCCG 9915
QY 538 GATGAAGAGTGAGCTCCCTTGGCCCTCAAGGAGCCCAAGGTGGGAGCACTCTTC 597
DB 9916 GATGAAGAGTGAGCTCCCTTGGCCCTCAAGGAGCCCAAGGTGGGAGCACTCTTC 9975
QY 598 C-GGGAGCAAGGCGCGCGCTGCGGAGCACTTCTCAGGCGAGACAGATTATTTCCGCT 656
DB 9976 CGGGAGCAAGGCGCGCGCTGCGGAGCACTTCTCAGGCGAGACAGATTATTTCCGCT 10035
QY 657 GCGGATCTGTGCTCCGAC-----CAGTTGTGTGCTCATCTGGAAGAGAGGCTT 710
DB 10036 GCGGATCTGTGCTCCGACAGTTTGTGTGCTCATCTGGAAGAGAGGCTT 10095
QY 711 GACCATTAAGAAATCATCTAAGAGACCGAGTCCCGGTAGATTCATGAAAAGAGAA 770
DB 10096 GACCATTAAGAAATCATCTAAGAGACCGAGTCCCGGTAGATTCATGAAAAGAGAA 10155
```

```
QY 771 CTCTGAGCTGACAGAGAGCGCTGTACCATTCATGCAACCCCAAGAGGAGCTTCTGAAGC 830
DB 10156 CTCTGAGCTGACAGAGAGCGCTGTACCATTCATGCAACCCCAAGAGGAGCTTCTGAAGC 10215
QY 831 ATGCCGATGATTTCTTGAATTCATGCAAGAGGCGAGATGAGCAACAACTAGCGGAGAA 890
DB 10216 ATGCCGATGATTTCTTGAATTCATGCAAGAGGCGAGATGAGCAACAACTAGCGGAGAA 10275
QY 891 GATTCCTGTGAAAATCTTGGGACACAAATGCTTGGTGGAAAGTATTTGAAAAGAG 950
DB 10276 GATTCCTGTGAAAATCTTGGGACACAAATGCTTGGTGGAAAGTATTTGAAAAGAG 10323
QY 951 CAGAAATTTGAAAATTTGAAACATGAAACAGGAGCCAAAGATACATCTCTTTCGA 1010
DB 10324 CAGAAATTTGAAAATTTGAAACATGAAACAGGAGCCAAAGATACATCTCTTTCGA 10383
QY 1011 GATTTTGGCATATACACCGGAAAGAACATCATCTGTAAGGAGCAG--TTGAGGCT 1069
DB 10384 GATTTTGGCATATATACACCGGAAAGAACATCATCTGTAAGGAGCAGTTTGAAGGCT 10443
QY 1070 GTGCCAGTGTGAGATGAGATTAAGAGAAAGCTGCTGAGGCTTTGAAAATGATATGC 1129
DB 10444 GTGCCAGTGTGAGATGAGATTAAGAGAAAGCTGCTGAGACCTTTGAAAATGATATGC 10503
QY 1130 TGGCTGTAAACAACCACTCCGATCTTCTCCAGCTGTACCCCATCAACGATTGGCC 1189
DB 10504 TGGCTGTAAATACGACCTCCGATCTTCTCCAGCTGTACCCCATCAACGATTGGCC 10563
QY 1190 CGTCCCGCATATACCTTTATCCAGAGCAGAGATTGTGATCTCTTATCCCAACC 1249
DB 10564 CGTCCCGCATATACCTTTATCCAGAGCAGAGATTGTGATCTCTTATCCCAACC 10623
QY 1250 AAGCTGTGGCGCATCATCGGAAAGAGGAGGACACATCAACACGCTGGAGATTGG 1309
DB 10624 AAGCTGTGGCGCATCATCGGAAAGAGGAGGACACATCAACACGCTGGAGATTGG 10683
QY 1310 CCGAGCTCTATCAAGATGCCCCCTGCGGAGAGCCAGACGTCAGCGAAAGATGTGCA 1369
DB 10684 CCGAGCTCTATCAAGATGCCCCCTGCGGAGAGCCAGACGTCAGCGAAAGATGTGCA 10743
QY 1370 TCATCAACCGGAGCCACCGAAGGCCAGTTCAAGGCCAGGAGCGGATCTTGGGAACTGA 1429
DB 10744 TCATCAACCGGAGCCACCGAAGGCCAGTTCAAGGCCAGGAGCGGATCTTGGGAACTGA 10803
QY 1430 AAGAGGAAAATCTTTTAAACCCCAAGAAAGATGAAGCTGGAAGCATATCAAGTGC 1489
DB 10804 AAGAGGAAAATCTTTTAAACCCCAAGAAAGATGAAGCTGGAAGCATATCAAGTGC 10863
QY 1490 CCTCTTCCACAGCTGCGCGGATGATGCGAAGAGTGGCAACCGTGAACGACTGAGA 1549
DB 10864 CCTCTTCCACAGCTGCGCGGATGATGCGAAGAGTGGCAACCGTGAACGACTGAGA 10923
QY 1550 ACTTAACAGTGCAAGATCATCTGCTCTGTGACCAAAACGCCAGATGAATAAGAGAG 1609
DB 10924 ACTTAACAGTGCAAGATCATCTGCTCTGTGACCAAAACGCCAGATGAATAAGAGAG 10983
QY 1610 TGAATGCAAAATTAATGCGGCACTTCTTGTGACCAAGCTGCAACAGCGGAAATCAGGG 1669
DB 10984 TGAATGCAAAATTAATGCGGCACTTCTTGTGACCAAGCTGCAACAGCGGAAATCAGGG 11043
QY 1670 AAATTTGACAAAGGTGAAGACCAAGAGCAAGAAATTAACCTGAGGAGTGCCTCAGAG 1729
DB 11044 AAATTTGACAAAGGTGAAGACCAAGAGCAAGAAATTAACCTGAGGAGTGCCTCAGAG 11103
QY 1730 GAGCAAGTAGGCTCCACAGGACACAGCAAAACAGATGATGATGATGATGATGATGATG 1789
DB 11104 GAGCAAGTAGGCTCCACAGGACACAGCAAAACAGATGATGATGATGATGATGATGATG 11163
QY 1790 ACCTGACAGATGAGACCAACAGGACGACGATGAGGAGCAACCAAGACCAATCTG 1849
DB 11164 ACCTGACAGATGAGACCAACAGGACGACGATGAGGAGCAACCAAGACCAATCTG 11223
QY 1850 AGGATGAGAGTCTGGGAGGCGGCGAGGAGCTCTGCGAGGCGCTGAGAAACCCAGAGG 1909
```

Db 11224 AGAATGAGAAAGCTGCGAGAGCTGCGACGAGACTCGAAGAGAGCCCTGGAATTCACAGG 11263
QY 1910 GCCGAGAGAGGGCGGGGAGAGTCAAGCAGATTGTCAGAAACGACGAGCC -CGCTCCC 1968
Db 11284 GCCGAGAGAGGGGTGGGAAAGTCAAGCAGATTGTCAGAAACGACGAGCCCTGTCAC 11343
QY 1969 GCGCCCGAGGGCTTTCGAGAGCTTCAAGCAGATTGTCAGAAACGACGAGCCCTGTCAC 2028
Db 11344 CCGCATCAGGCTTTCGAGAGCTTCAAGCAGATTGTCAGAAACGACGAGCCCTGTCAC 11403
QY 2029 GAGCTCCAGAGCGCTATCCCTTTTGTGAACTCAACATGAGTGAACGTTCAAGCCA 2088
Db 11404 TAACGCCACAGCGCTATCCCTTTTGTGAACTCAACATGAGTGAACGTTCAAGCCA 11461
QY 2089 AGCAAAATGACACACCTTTTCTGTGCAAAATCGTCTGTATCATGTGTATCATATTAG 2148
Db 11462 AGCAAAATGACACACCTTTTCTGTGCAAAATCGTCTGTATCATGTGTATCATATTAG 11521
QY 2149 AAAAGGAAGATGTTAAGATATGTGGCTGTGGTTACACAGGGCTCTGACAGCGTAAAT 2208
Db 11522 AAAGGAAGATGTTAAGATATGTGGCTGTGGTTACACAGGGCTCTGACAGCGTAAAT 11581
QY 2209 TATTTTGAATATATATATCAATTAATCTCAACTCACTCAATTTTAACTAATTAAT 2268
Db 11582 TATTTTGAATATATATATCAATTAATCTCAACTCACTCAATTTTAACTAATTAAT 11636
QY 2269 TTTTCTCTTTTAAAGAGAGAGCGCTTTCTAGACTTAAGATTAATTAATTAATTAAT 2323
Db 11637 TTTTCTCTTTTAAAGAGAGAGCGCTTTCTAGACTTAAGATTAATTAATTAATTAAT 11696
QY 2324 TTTTGGAGGTCTCAAGGTATGAGAGAGAGCTTTGAGGCGACCGGACAAATTAATCAAC 2383
Db 11697 TTTTGGAGGTCTCAAGGTATGAGAGAGAGCTTTGAGGCGACCGGACAAATTAATCAAC 11756
QY 2384 AGGGAATCTGTGTGAGAGAGAGAGCTCAAGAGTTCAGATCACTGTGTATGATCAAG 2443
Db 11757 AGGGAATCTGTGTGAGAGAGAGAGCTCAAGAGTTCAGATCACTGTGTATGATCAAG 11816
QY 2444 AAGGATACCGTCTCTGAGAGAGAGAGCTGTATCACTCTCAAGCTGTGTATGATCAAG 2503
Db 11817 AAGGATACCGTCTCTGAGAGAGAGAGCTGTATCACTCTCAAGCTGTGTATGATCAAG 11876
QY 2504 CACCATTTCTCTGTGCTCAAGAGTCTTAACTGTTTGTGATCACTGTGTATGATCAAG 2563
Db 11877 CACCATTTCTCTGTGCTCAAGAGTCTTAACTGTTTGTGATCACTGTGTATGATCAAG 11936
QY 2564 CTCTGTCTCTGTGTATGATCACTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2623
Db 11937 ATCATCTCTCTGTGTATGATCACTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11996
QY 2624 CTTTGAATTTCTGATCCCTCACTCTCACTCTCACTCTCACTCTCTCTCTCTCTCTCTCTCT 2682
Db 11997 CTTTGAATTTCTGATCCCTCACTCTCACTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCT 12056
QY 2683 GGGAAAGCAGTGTCTGATGATCACTCAACAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2742
Db 12057 GGGAAAGCAGTGTCTGATGATCACTCAACAAAGAAAGAAAGAAAGAAAGAAAGAAAG 12116
QY 2743 CAGCTTCACTTCACTGTGTTACTCAAAAGAAAGAAAGTCAATGTGTTGTTCTTCAAG 2802
Db 12117 CAGCTTCACTTCACTGTGTTACTCAAAAGAAAGAAAGTCAATGTGTTGTTCTTCAAG 12176
QY 2803 TTTTGAAGAGGAAAG 2862
Db 12177 TTTTGAAGAGGAAAG 12236
QY 2863 TCCACATGAAAGATGATTTTGTGCT - TTTTGAATTTTGTGTTAAGCATCAATAT 2921
Db 12237 TCCACATGAAAGATGATTTTGTGCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 12296
QY 2922 CAGCAAAATGATCTTCTTTTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2981

Db 12297 CAGCAAAATGATTTCTTCTTTT - AAAAAAATGTGAGAGAGAGAGAGAGAGAGAG 12353
QY 2982 GTTGTGGCGCCAGGCGGTTAAATTAATCAAGATTTTAAATGAGAGAGAGAGAGAGAG 3041
Db 12354 GTTGTGGCGCCAGGCGGTTAAATTAATTAATGAGATTTTAAATGAGAGAGAGAGAG 12413
QY 3042 GCTACCTCAGGTG - TTTTACCTCAGCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 3100
Db 12414 GCTGCTCAGGAGGT 12473
QY 3101 AAAGCTGATGATGAGAGATTTTCTTTT - TTTTAAATTAATTAATGAGAGAGAGAGAG 3159
Db 12474 AAAGCTGATGATGAGAGATTTTCTTTT - TTTTAAATTAATTAATGAGAGAGAGAG 12533
QY 3160 AAAGATATCACTCCAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3219
Db 12534 AAAGATATCACTCCAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 12593
QY 3220 TTTGCTTCCGCTAGGCAAG 3279
Db 12594 TTTGCTTCCGCTAGGCAAG 12653
QY 3280 ATTT 3283
Db 12654 ATTT 12657
RESULT 11
AC015706 182695 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 1 clone RP11-325P15 map 1, WORKING DRAFT
DEFINITION
SEQUENCE, 17 unordered pieces.
AC015706
AC015706.3 GI:8096833
VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 182695)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouckhalter,B.,
Brown,A., Castle,A., Collins,S., Colangelo,M., Collins,S., Colangelo,M.,
Cooke,P., DeArillano,K., Dewar,K., Domingo,M., Donnell,J., Doyle,M.,
Fitzgerald,P., Fitzgerald,W., Forrester,C., Funke,R., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heathford,A., Horton,L., Howland,J.,
Hoczky,J., Liew,C., Locke,K., MacDonald,P., Margolis,A., Klein,J.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Melnick,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Poltara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Homann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teste,S., Tittell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 182695)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bedalov,F.,
Boguslavsky,L., Bouckhalter,B., Brown,A., Buckett,G., Campiano,A.,
Castle,A., Choquel,Y., Colangelo,M., Collins,S., Colangelo,M.,
Cooke,P., DeArillano,K., Dewar,K., Diaz,J., Dodge,S., Domingo,M.,
Doyle,M., Ferreira,P., Fitzgerald,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B.,
Heathford,A., Horton,L., Howland,J., Hoczky,J., Iliiev,I., Johnson,R.,
Jones,C., Kann,L., Karatas,A., Klein,J., Lehoucq,K., Lamazares,R.,
Lander,E., Lehoucq,J., Levine,R., Liew,C., Liu,G., Locke,K., MacDonald,P., Margolis,N.,

TITLE
JOURNAL
COMMENT

McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheters, R.,
Meldrum, J., Meneus, J., Milnova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Teefaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, M.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 26, 2000 this sequence version replaced gi:6454072.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L1382

Center clone name: 325 P.15

Summary Statistics

Sequencing vector: M13; W77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960721

Consensus quality: 141788 bases at least Q40

Consensus quality: 163541 bases at least Q30

Consensus quality: 174666 bases at least Q20

Insert size: 181000; agarose-ftp

Insert size: 181095; sum-of-contigs

Quality coverage: 4.2 in Q20 bases; sum-of-contigs

Quality coverage: 4.2 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently

consists of 17 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

```

1      2035: contig of 2035 bp in length
*      2036      2135: gap of 100 bp
*      2037      6069: contig of 3934 bp in length
*      6070      6169: gap of 100 bp
*      6170      10460: contig of 4251 bp in length
*      10461      10560: gap of 100 bp
*      10561      14300: contig of 3740 bp in length
*      14301      14400: gap of 100 bp
*      14401      18480: contig of 4080 bp in length
*      18481      18580: gap of 100 bp
*      18581      23093: contig of 4513 bp in length
*      23094      23193: gap of 100 bp
*      23194      28213: contig of 5020 bp in length
*      28214      32808: contig of 4495 bp in length
*      32809      32908: gap of 100 bp
*      32909      37000: contig of 4052 bp in length
*      37001      37100: gap of 100 bp
*      37101      45183: contig of 8083 bp in length
*      45184      45283: gap of 100 bp
*      45284      53421: contig of 8138 bp in length
*      53422      53521: gap of 100 bp
*      53522      64476: contig of 10955 bp in length
*      64477      64576: gap of 100 bp
*      64577      75144: contig of 10568 bp in length
*      75145      75244: gap of 100 bp
*      75245      93230: contig of 17986 bp in length
*      93231      93330: gap of 100 bp

```

FEATURES
source
* 93331 116951: contig of 23621 bp in length
* 116952 117051: gap of 100 bp
* 117052 143611: contig of 28560 bp in length
* 143612 143711: gap of 100 bp
* 143712 182695: contig of 38984 bp in length.
Location/Qualifiers
1..182695
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-325P15"
/clone_id="RP11-11 Human Male BAC"
1..2035
/note="assembly_fragment"
2136..6069
/note="assembly_fragment"
6170..10460
/note="assembly_fragment"
10561..14300
/note="assembly_fragment"
14401..18480
/note="assembly_fragment"
18581..23093
/note="assembly_fragment"
23194..28213
/note="assembly_fragment"
28314..32808
/note="assembly_fragment"
clone_end:17
vector_side:right
32909..37000
/note="assembly_fragment"
37101..45183
/note="assembly_fragment"
45284..53421
/note="assembly_fragment"
clone_end:SP6
vector_side:left
53522..64476
/note="assembly_fragment"
64577..75144
/note="assembly_fragment"
75245..93230
/note="assembly_fragment"
93331..116951
/note="assembly_fragment"
117052..143611
/note="assembly_fragment"
143712..182695
/note="assembly_fragment"

ORIGIN

Query Match 84.7%; Score 2779.8; DB 2: Length 182695;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 3088; Conservative 0; Mismatches 168; Indels 48; Gaps 16;

```

QY      1  GGGAGGAGGAGGAGGAGGAGGCGCGGATACGGGCGCGGAGAGCGCGGGCTCTCGGGG 60
DB      19246  GGGAGGAGGAGGAGGAGGAGGCGCGGATACGGGCGCGGAGAGCGCGGGCTCTCGGGG 19305
QY      61  AAGAGACGATGATGAACAAGCTTTATCATCGGAACTGAAGCCCGCGTACCGCCGAC 120
DB      19306  AAGAGACGATGATGAACAAGCTTTATCATCGGAACTTAAGCCCTGCGGTACCGCCGAC 19365
QY      121  GACCTCGGAGCTCTTTGGGGAACAGAACTGCTCCCTGGCGGACAGGTCTGCTGAAG 180
DB      19366  GACCTCGGAGCTCTTTGGGGAACAGAACTGCTCCCTGGCGGACAGGTCTGCTGAAG 19425
QY      181  TCGGCGTACCGCTTCGCGAATACCGGACAGAACTGGGCAATCCCGCATCGAGACC 240
DB      19426  TCGGCGTACCGCTTCGCGAATACCGGACAGAACTGGGCAATCCCGCATCGAGACC 19485

```

QY 241 CTCTGGGTAAAGTGAATTCGATGGAAAAATCATGGAAGTTGATTACTCAGTCTCTAA 300
|
Db 19486 CTCTGGGTAAAGTGAATTCGATGGAAAAATCATGGAAGTTGATTACTCAGTCTCTAA 19545
QY 301 AAGCTTAAGGAGCAGAAAAATTCAGATTCGAAACATCCCTCTCAGCTGAGTGGAGGTG 360
|
Db 19546 AAGCTTAAGGAGCAGAAAAATTCAGATTCGAAACATCCCTCTCAGCTGAGTGGAGGTG 19605
QY 361 TTGATGAGCTTTTGGCTCAATATGGAACAGTGGGAATGTGG---AACAGTCAACA 417
|
Db 19606 TTGATGAGCTTTTGGCTCAATATGGAACAGTGGGAATGTGGAAACAACTCAACA 19665
QY 418 GACACAGAAACCGCGTGTCAACGTCACATATGCAACAGAGAAAGAAATAATAC 477
|
Db 19666 GACACAGAAACCGCTGTGTCAACGTCACATATGCAACAGAGAAAG---AAATTAAC 19720
QY 478 ATGAGAAAGCTPAAGCGGAGATCAAGTTGAGAACTACTCTTCAAGATTTCTACATCCG 537
|
Db 19721 ATGAGAGGCTPAAGCGGAGATCAAGTTGAGAACTAGTCTTCAAGATTTCTACATCCG 19780
QY 538 GATGAAGAGTGAAGTCCCTTCCCTTCCCTCAGGAGAGCCAGCGTGGGAGCACTCTCC 597
|
Db 19781 GATGAAGAGTGAAGTCCCTTCCCTTCCCTCAGGAGAGCCAGCGTGGGAGCACTCTCC 19840
QY 598 C-GGGAGAGAGGAGCAGCGCCCTGGGGGCACTTCTCAGGAGCAAGAGATGATTTCCGCT 656
|
Db 19841 CGGGAGAGAGAGCAGCGCCCTGGGGGCGCTTCTTGAAGCAGAGATGATTTCTCTCT 19900
QY 657 GCGGATCTGTGCTCCAC-----CAGTTTGTGTGTCATCATCGAAAGAGGCTT 710
|
Db 19901 GCGAGTCTGTGCTCCCTCAGTTTGAGTTGTTGTGTCTCATCATCAAAAGAGGCTT 19960
QY 711 GACCAATAAGAACTACCTAGCAGAGCCCACTCCCGGTAAATTCATAGAAAGAGAA 770
|
Db 19961 GACCAATAAGAACTACCTAGCAGAGCCCACTCCCGGTAAATTCATAGAAAGAGAA 20020
QY 771 CTCTGAGCTGCAAGAACTGTCAACCATCATGCGCAAGGAGGAGCTTCTGAGC 830
|
Db 20021 CTCTGAGCTGCAAGAACTGTCAACCATCATGCGCAAGGAGGAGCTTCTGAGC 20080
QY 831 ATGCGCATGATTTCTTGAATCATGCAAGAAAGAGCAGATGAGACCAATGAGCCGAGA 890
|
Db 20081 ATGCGCATGATTTCTTGAATCATGCAAGAAAGAGCAGATGAGACCAATGAGCCGAGA 20140
QY 891 GATTCCCTGAAAAATCTTGGCAACAATGCTTGTGTGAAAGACTGATGAAAAAGAG 950
|
Db 20141 GATTCCCTGAAAAATCTTGGCCCAATGCTGTGTT-----GGAAGAGAG 20188
QY 951 CAGAAATTTGAAGAAAAATGAAATGAAACAGGAGCAAGATACATTCATCTTGA 1010
|
Db 20189 CAGAAATTTGAAGAAAAATGAAATGAAACAGGAGCAAGATACATTCATCTTGA 20248
QY 1011 GGAATTGAGCATATACAAACCGGAAAGAACCATCATGCTGAGAGGCAAG-TTGAAGCT 1069
|
Db 20249 GGAATTGAGCATATATATACAGGAAAGATCATCATGTAAGGCAAGTTGAGGCT 20308
QY 1070 GTGCCATGCTGAGATGAGATTTGAAGAGCTGCTGAGGCTTTGAAATGATATGC 1129
|
Db 20309 GTGCCATGCTGAGATGAGATTTGAAGAGCTGCTGAGGCTTTGAAATGATATGT 20368
QY 1130 TGGCTGTTAACACCACTCCGATACTTCTCAGGCTGTACCCCATCACAGTTGGCC 1189
|
Db 20369 TGGCTGTTAATAGCACTCCGATACTTCTCAGGCTGTACCCCATCACAGTTGGCC 20428
QY 1190 CGTTCCGCAATCACTCTTATCAGAGCAGAGATGTAATCTTTCATCCCAACC 1249
|
Db 20429 CGTTCCGCAATCACTCTTATCAGAGCAGAGATGTAATCTTTCATCCCAACC 20488
QY 1250 AGGCTGAGGAGCAGATGAGGAGAAAGGAGGAGCAATCAACAGCTGGCGAGATTGC 1309
|
Db 20489 AGGCTGAGGAGCAGATGAGGAGAAAGGAGGAGCAATCAACAGCTGGCGAGATTGC 20548

QY 1310 CCGAGCCTTATCAAGATTGCCCTGGGAAAGGCCAGAGCTACGCCAAAGAGTGTCA 1369
|
Db 20549 CCAAGCCTTATCAAGATTGCCCTGGGAAAGGCCAGAGCTTCAACAAAGAGTGTCA 20608
QY 1370 TCATACCGGAGCAGCCGAAAGGCCAGTTCAGAGGCCAGAGAGCGATCTTGGGAACTGA 1429
|
Db 20609 TCATACCGGAGCAGCCGAAAGGCCAGTTCAGAGGCCAGAGAGCGATCTTGGGAACTGA 20668
QY 1430 AAGAGAAAACTTCTTAAACCCCAAGAAAGTGAAGCTGAGAGCGATATCAAGTGC 1489
|
Db 20669 AAGAGAAAACTTCTTAAACCCCAAGAAAGTGAAGCTGAGAGCGGTATCAAGTGC 20728
QY 1490 CCTCTTCAACAGCTGCGGAGTGTGGCAAAAGTGGCAAGACCTGTAAACGACTGCA 1549
|
Db 20729 CCTCTTCAACAGCTGCGGAGTGTGGCAAAAGTGGCAAACTGTGAATGAATGCA 20788
QY 1550 ACTTAACAGTGCAGAAAGTCAATGCTGCGAGCAAAAGCAGATGAAATGAGGAAG 1609
|
Db 20789 ACTTAACAGTGCAGAAAGTCAATGCTGCGAGCAAAAGCAGATGAAATGAGGAAG 20848
QY 1610 TGATGTCAGAAATTATCGGAGCACTTCTTGTAGCCAGACTGCAAGGCAAGATCAAGG 1669
|
Db 20849 TGATGTCAGAAATTATGAGCACTTCTTGTAGCCAGACTGCAAGGCAAGATCAAGG 20908
QY 1670 AAATTGTACACAGGTGAAAGAGCAGAGAGCAAAATACCTCAGGAGTCCCTTCAAGC 1729
|
Db 20909 AAATTGTACACAGGTGAAAGAGCAGAGAGCAAAATACCTCAGGAGTCCCTTCAAGC 20968
QY 1730 GCAGCAAGTGAAGCTCCCAAGGAGCAGCAAGCAAAACAGATGATGAGGCTTCCAG 1789
|
Db 20969 ACAGCAAGTGAAGCTCCCAAGGAGCAGCAAGCAAAACAGATGATGAGGCTTCCAG 21028
QY 1790 ACTGACAGATGAGACCAACGAGCCAGATCGGAGCAAAACCAAGATGATGAGGCTTCCAG 1849
|
Db 21029 ACTGACAGATGAGACCAACGAGCCAGATCGGAGCAAAACCAAGATGATGAGGCTTCCAG 21088
QY 1850 AGGATGAGAAAGTGTGGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1909
|
Db 21089 AGGATGAGAAAGTGTGGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 21148
QY 1910 GCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1968
|
Db 21149 GCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 21208
QY 1969 GCCCCCAAGGCTTCTGAGGCTTCAAGCCATCACTTCAACATCACTCGGATCTCTCT 2028
|
Db 21209 CCCCATCAGGCTTCTGAGGCTTCAAGCCATCACTTCAACATCACTCGGATCTCTCTCT 21268
QY 2029 GAATCCCAAGGAGCTATCCCTTTTGTAGTGAATCAATGAGTGAAGTGTTCAGAGCA 2088
|
Db 21269 TAACTCCCAAGGAGCTATCCCTTTTGTAGTGAATCAATGAGTGAAGTGTTCAGAGCA 21326
QY 2089 AGCAAAATGACACCCCTTTTCTGTGCAATGCTCTGTGAACATGAGTGAATATAG 2148
|
Db 21327 AGCAAAATGACACCCCTTTTCTGTGCAATGCTCTGTGAACATGAGTGAATATAG 21386
QY 2149 AAAAGGAGAGTGTAAAGATATGAGGCTGTGAGTTACAGAGGCTCTGAGCGGTATA 2208
|
Db 21387 AAAAGGAGAGTGTAAAGATATGAGGCTGTGAGTTACAGAGGCTCTGAGCGGTATA 21446
QY 2209 TATTTTGAATTAATATATCAATTAATCAATTAATCAATTAATCAATTAAT 2268
|
Db 21447 TATTTTGAATTAATATATCAATTAATCAATTAATCAATTAATCAATTAAT 21501
QY 2269 TTTTCTTTCTTTTAAAGAGAAAGCGGCTTTTCAAGCTTTTAAAGATTA 2323
|
Db 21502 TTTTCTTTCTTTTAAAGAGAAAGCGGCTTTTCTAGACTTTTAAAGATTAAGTCTTTTC 21561
QY 2324 TTTGGAGGCTTCAAGGCTGTAGAGAGGCTTTGAGGCAACCCGCAAAATTCACCCAG 2383
|
Db 21562 TTTGGAGGCTTCAAGGCTGTAGAGAGGCTTTGAGGCAACCCGCAAAATTCACCCAG 21621
QY 2384 AGGGAATCTGCTGAGAGAGCACTCAAGGAGTCTGATGACTGTGTATGTCAGAG 2443

Dp	21622	AGGGAATCTCAAAAGGAAGAACATCAAGCAGTTCTGGATCACTGTGTATGTCAACAG	21681
Qy	2444	AAGGATACCGTCTCCTTGAAGAAGAACTCTGTCACTCTCATGGCTGTACGTCATA	2503
Dp	21682	AAGGGATACCATCTCCTTGAAGAAGAACTCTGTCACTCTCATGGCTGTACGTCATA	21741
Qy	2504	CACCCATTTCTCTTTCCTTCAACAGTTTTAACTGGTTTTTTTGGCATACGCTATAAAT	2563
Dp	21742	CACCCATTTCTCTTTCCTTCAACAGTTTTAACTGGTTTTTTTGGCATACGCTATAAAT	21801
Qy	2564	CTCTGTCTCTCTGTGTTTATCTCTCCCTCCCTCCCTCCCTCTTCTTCATCTCCAT	2623
Dp	21802	ATCCATCTCTCTGTGTTTATCTCTCCCTCCCTCCCTCCCTCTTCTTCATCTCCAT	21861
Qy	2624	CTTTTGAATTCGTCATCCCTCCATCTCAATCCCGATCTACGACCC-CCCCCCCCCA	2682
Dp	21862	CTTTTGAATTCGTCATCCCTCCATCTCAATCCCGATCTACGACCC-CCCCCCCCCA	21922
Qy	2683	GGCAAAAGCAGTGTCTGTAGTATCAATCACACAAAGAAACAAAGCGAAACACACAAAC	2742
Dp	21922	GGCAAAAGCAGTGTCTGTAGTATCGATCAACACAAAGAAACAAAGGTGAACACACAAAC	21981
Qy	2743	CAGCCTTAACCTTTCACCTTGGTAACTCAAAAGAACAGAGTCATGGATCTTGTCCTAGG	2802
Dp	21982	CAACCTCAATTTACACTGGTATCTCAAAAGAACAGAGTCATGGATGTATGTCTTAGG	2204
Qy	2803	TTTTTGAAGAGAAACAGAACCCACCAACCAACCAATCAACCAACCAAGAAAAAT	2862
Dp	22042	TTTTTGAAGAGAAACAGAACCCATCAACCAACCAATCAACCAACCAAGAAAAAT	22101
Qy	2863	TCACCAATGAAAGATGATTTTGTCT-TTTTTGATTTTGGTGTATTAAGCCATCATTT	2921
Dp	22102	TCACCAATGAAAGAGGTGTGTCTTTTGGATTTGGTGTATTAAGCCATCATTAAT	22161
Qy	2922	CAGCAAAATGATTCCTTTCTTAAAAAATAATGTGAGAAAGTATGAATTTTACCAAG	2981
Dp	22162	CAGCAAAATGATTTCTTTCTTT---AAAAAATGTGAGAAAGTATGAATTTTACCAAG	2221
Qy	2982	GTTGTGTGGCCAGGGCGTAAATTCACAGATTTTTTTTAAACAGAAAAACACAGAAAG	3041
Dp	22219	GTTGTGTGGCCAGGGCATTTAAATTTTAAAGATTTTTTTTAAAGGAAAAACACAGAAAA	2227
Qy	3042	GCTACCTCAGGTG-TTTTTAACCTAGACCTGTCTGTGTTCCCTTAGATTTTGT	3100
Dp	22279	GCTCCTCAGGTGTTTTTTTAACTAGACCTGTCTGTGTTCCCTTAGATTTTGT	2233
Qy	3101	AAAGCTGATAGTTGAGCATTTTTTTA-TTTTTTAAATAAAATGAGTGGAAAAAAT	3159
Dp	22339	AAAGCTGATAGTTGAGCATTTTTTTAATTTTTTAAATAAAATTAGTGGAAAAAAT	2239
Qy	3160	AAGATATCAATCGCCAGCTCTGGAAGAGTGCACATCCAAAGTGCCACAGTGTCTGAA	3219
Dp	22399	AAGATATCAATCGCCAGCTCTGGAAGAGTGCACATCCAAAGTGTCCACAGTGTCTGAA	2245
Qy	3220	TTGTCTTCGCGTAGCGCAAGACNATATGCGCTCTTTTGGACAAACCTTGAATGTT	3279
Dp	22459	TTGTCTTCGCGTAGCGCAAGACCTATATGCGCTCTTTTGGACAAACCTTGAATGTT	2251
Qy	3280	ATTT 3283	
Dp	22519	ATTT 22522	
RESULT 12			
AC104980			
LOCUS			
DEFINITION Homo sapiens chromosome 8, clone Rpl1-2b1D17, complete sequence.			
ACCESSION			
VERSION			
KEYWORDS			
HTG.			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
187226 bp DNA linear PRI 31-MAR-2002			

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 187226)
Birren, B., Linton, L., Nussbaum, C. and Lander, E.
Homo sapiens chromosome 8, clone RP11-281D17
Unpublished

2 (bases 1 to 187226)
Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campiorgio, A., Chang, U., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., D'arellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Gardes, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K.,
Lamarcares, R., Landers, T., Lepocky, J., Levine, R., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margus, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
Menelaus, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C.,
Nobdu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Pinnukhant, P., Pomeroy, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rossetti, M., Roy, A., Santos, R., Schauer, S., Schumpack, R.,
Semman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, K., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tophan, K., Travins, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome
Research 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187226)
Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campiorgio, A., Chang, U., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., D'arellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Gardes, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K.,
Lamarcares, R., Landers, T., Lepocky, J., Levine, R., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margus, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
Menelaus, L., Mihova, T., Mlenka, V., Murphy, T., Naylor, J., Nguyen, C.,
Nobdu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Pinnukhant, P., Pomeroy, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rossetti, M., Roy, A., Santos, R., Schauer, S., Schumpack, R.,
Semman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, K., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tophan, K., Travins, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia, Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 187226)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 8, clone RPL1-281D17
 Unpublished
 2 (bases 1 to 187226)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
 Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,
 Chopel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., Dearellano,K., Dewar,K., Diaz,U.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Girde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Jones,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
 Lamasers,R., Landers,T., Lebockzy,J., Levine,R., Liu,G.,
 Maclean,C., MacDonald,P., Major,J., Margus,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrim,J.,
 Menus,L., Mihova,T., Mieng,V., Murphy,T., Naylor,T., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunhhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,U., Rosetti,M., Roy,A., Santos R., Schaner,S., Schnapack,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,C.N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 187226)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
 Boukhalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J.,
 Chazaro,B., Chopel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,U.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Girde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Ginde,S., Horton,L., Horton,L., Iliev,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamasers,R.,
 Landers,T., Lebockzy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 Maclean,C., MacDonald,P., Major,J., Margus,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Menus,L.,
 Mihova,T., Mieng,V., Murphy,T., Naylor,T., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunhhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos R., Schaner,S., Schnapack,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 187226)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
 Boukhalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J.,
 Chazaro,B., Chopel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,U.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Girde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamasers,R.,
 Landers,T., Lebockzy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 Maclean,C., MacDonald,P., Major,J., Margus,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Menus,L.,
 Mihova,T., Mieng,V., Murphy,T., Naylor,T., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunhhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos R., Schaner,S., Schnapack,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

TITLE
JOURNAL

COMMENT

Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rie, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schipack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Struss, N., Subramanian, A., Talamas, U., Testaye, S., Theodore, J., Topham, K., Travers, M., Trivis, N., Trigilio, J., Vassiliev, H., Viet, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. T., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (31-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 31, 2002 this sequence version replaced gi:19683374.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: 122878
Center clone name: 281_D_17

FEATURES
source

----- Location/Qualifiers
1..187226
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-281D17"
/clone_lib="RPCT-11 Human Male BAC"
1..5915
/rpt_family="L1PA5"
5916..5943
/rpt_family="AT_rich"
5944..6477
/rpt_family="L1"
6563..8886
/rpt_family="MER2"
repeat_region
complement(9296..9555)
/rpt_family="L2"
repeat_region
complement(9865..10094)
/rpt_family="MIR"
repeat_region
complement(10346..10504)
/rpt_family="AluSg/x"
repeat_region
complement(10821..11111)
/rpt_family="AluSc"
11162..11300
/rpt_family="MER47A"
11301..11596
/rpt_family="AluSx"
repeat_region
11597..11836
/rpt_family="MER47A"
12185..12319
/rpt_family="FLAM_C"
12520..12822
/rpt_family="AluSp"
repeat_region
complement(12885..13167)
/rpt_family="AluYb"
repeat_region
complement(13230..13379)
/rpt_family="MIR"
repeat_region
complement(14179..14479)
/rpt_family="AluSg"
repeat_region
complement(14630..14747)
/rpt_family="MIR"
14748..15055
/rpt_family="AluSq"
repeat_region
complement(15056..15145)
/rpt_family="MIR"
repeat_region
complement(15206..15400)
/rpt_family="MIR"

	repeat_region	15576..15606	/rpt_family="AT_rich"
	repeat_region	complement(15607..15901)	/rpt_family="AluSp"
	repeat_region	16127..16556	/rpt_family="L1ME4A"
	repeat_region	16657..16684	/rpt_family="AT_rich"
	repeat_region	complement(17149..17513)	/rpt_family="MLT1A1"
	repeat_region	17966..18067	/rpt_family="MER47B"
	repeat_region	complement(18068..18365)	/rpt_family="AluYo"
	repeat_region	18366..18666	/rpt_family="MER47B"
	repeat_region	19538..19561	/rpt_family="AT_rich"
	repeat_region	complement(19562..21070)	/rpt_family="L1PA5"
	repeat_region	complement(21072..21141)	/rpt_family="FLAM_C"
	repeat_region	22731..22771	/rpt_family="MIR"
	repeat_region	complement(23445..23902)	/rpt_family="L1PA5"
	repeat_region	complement(25000..25014)	/rpt_family="FLAM_C"
	repeat_region	25015..25083	/rpt_family="TTCA)n"
	repeat_region	complement(25084..25200)	/rpt_family="FLAM_C"
	repeat_region	25393..25450	/rpt_family="AT_rich"
	repeat_region	complement(25762..26547)	/rpt_family="L1ME4A"
	repeat_region	26823..26889	/rpt_family="TAGA)n"
	repeat_region	27373..27795	/rpt_family="LTR37A"
	repeat_region	complement(29177..29247)	/rpt_family="MIR"
	repeat_region	29248..29370	/rpt_family="AluYo"
	repeat_region	29371..29395	/rpt_family="CA)n"
	repeat_region	29396..29562	
Query Match		70.9%; Score 2327; DB 9; Length 187226;	
Best Local Similarity		89.4%; Pred. No. 0;	
Matches 2726; Conservative		0; Mismatches 211; Indels 113; Gaps 16;	
QY	247	GGTAAAGTGAATTCATGAGGAAATATCATGGAAGTTGATTACTAGTCTCTAAAGGCTA	306
Db	79640	CGTAAAGTGAATTCATGAGGAAATATCATGGAAGTTGATTACTAGTCTCTAAAGGCTA	79699
QY	307	AGAGACGAGAAATTCAGATTCGAAACATCCTCTCATCCTGAGTGGAGTGTGGAT	366
Db	79700	AGGAAACGAGAAATTCAGATTCGAAACATCCTCTCATCCTGAGTGGAGTGTGGAT	79759
QY	367	GGACTTTTGGCTCAATATGAGACAGTGGAGAAATGTGGAACAAGTCAACAGACAGAA	426
Db	79760	GCACCTTTGGCTCAACACGGGAAATGTGGAGAAATGTGGAACAAGTCAACAGACAGAA	79819
QY	427	ACGCGCGTGTCAACGTCATATGCAACAAGAGAAAGAAATAGCATGGAAG	486
Db	79820	ACTGCTGTGTCAATGTCATATGCAACAAGAGAAAGC-AAAATAGCCATGGAAG	79878
QY	487	CTAAGCGGAGCATGTTGAGAACTACTCTTCAAGATTTCTCAATCCCGATGAAG	546
Db	79879	CTAAGCGGAGCATGTTGAGAACTATTCCTTCAAGATTTCTCAATCCCGATGAAG	79938
QY	547	GTAGGCT-CCCTTCCGCCCTCAGGAGCCCAAGCGTGGGAGCACTTTCCCGGAGCA	605


```
Db 79939 GTGAGCTCCCCCTTTACCCCTCAGCGAGCCAGGCTGGGAGCACTCTT -CCGGAGCA 79997
Qy 606 AGGCCACGCCCTTGGGGCACTTCTCAGGCCACAGATTGATTTCCGCTGGCATCTT 665
Db 79998 AGGTACAG-CCCTGGGGGCTCTTCTCAGGCCGAGCAGACTGATTTCCGCTGGCATCCA 80056
Qy 666 GGTCCCCACCCAGTTTGTGTGTCATCATCGGAAAGAGGGCTTGACCATTAAGACAT 725
Db 80057 -----ACCGAGTTTGTGTGTCATGCTGAAAGAGGGCTTGACCATGTGACAT 80109
Qy 726 CACTAAGCAGACCCAGTCGCCGGGTAGATATCCATAGAAAGAAAGAACTGTGAGCTGCAG 785
Db 80110 CACTAGCAGACCCAGTCCCGGGTAGACATCATAGAAAGAAAGAACTGTGAGCTGCAG 80169
Qy 786 GAAAGCTGTCACTTCATCCATGCCACCCAGAGGGGAACTTGTAGAGCATGGCCGATGATTT 845
Db 80170 GAAAGCTGTCACTTCATCCATGCCACCCAGAGGGGAACTTGTAGAGCATGGCACAGATTTCT 80229
Qy 846 TGAATTCATGCAAGAAAGAGGAGATGAGACCAACTAGCCGAAAGATTCCTCTGAAAT 905
Db 80230 TGAAGACATGCAAGAAAGAGGAGATGAGACCAACTAGCCGAAAGATTCCTCTGAAAT 80289
Qy 906 CTGGCACAACATGGCTTGTGTGAAAGACTGATTTGAAAGAAAGAGCAAAATTTGAGAA 965
Db 80290 CGTGCCCAACATGGCTTGTGTGAAAGACTGATTTGAAAGAAAGAGCAAAATTTGAGAA 80349
Qy 966 AATTGAACATGAACAGGACCAAGATTAACATCTCATCTTTGCAAGATTTGAGCATATA 1025
Db 80350 AATGAACATGAACAGGACCAAGATTAACATCTCATCTTTGCAAGATTTGAGCATATAG 80409
Qy 1026 CAACCCGAAAGAACCATCACTGTGAAAGGCAACGTGAGGCGCTGGCCAGTGTAGAT 1085
Db 80410 CAACCCGAAAGAACCATCACTGTGAAAGGCAACGTGAGGCGCTTATGCCAGTGTAGAT 80469
Qy 1086 AGAGATTATGAAGAAAGCTGCTGAGGCTTTGAAATGATATGCTGCTTTAACCACA 1145
Db 80470 GAGGATTATGAAGAAAGCTGCTGAGGCTTTGAAATGATATGCTGCTTTAATACCA 80529
Qy 1146 CTCCGATATC-TTCTCAGGCTGTATACCCCATCAGATTTGGCCGCTCCGSCATCATC 1204
Db 80530 CTCCGATATCTTCTCAGGCTGTATACCCCATCAGATTTGGCCGCTCCGSCATCAG 80589
Qy 1205 ACTCTATCCAGAGCAGAGATTTGATGATCTTCAATCCCAACCCAGGCTGTGGGCGCA 1264
Db 80590 ATTCTTATCAGAGCAGAGATTTGATGATCTTCAATCCCAACCCAGGCTGTGGGATCA 80649
Qy 1265 TCATCGGAAAGAAAGGGGCGACATCATCAACAGCTGGAGATTTGCTCGAGCTTATCA 1324
Db 80650 TCATCGGAAAGAAAGGGGCGACATCATCAACAGCTGGAGATTTGCTCGAGCTTCA--- 80706
Qy 1325 AGATTGCCCTGCGGAGGCCCAAGCTCAGCGAAAGATGTATCATCATCCGGGCGAC 1384
Db 80707 ----- 80706
Qy 1385 CGGAGGCCAGTTCAAGGCCAGGAGCGATCTTTGGGAAATGAAAGAGAAATCTTCT 1444
Db 80707 -----TCAGGCCAGGAGCGATCTTCCGAAATCTTAAAGAAAGAAATCTTCT 80754
Qy 1445 TTAAACCCAAAGAAAGATGAGCTGAGCGATATCAGATGCTCTTCCAGAGCTG 1504
Db 80755 TTAAACCCAAAGAAAGATGAGCTGAGCGATATCAGATGCTCTTCCAGAGCTG 80814
Qy 1505 GCGGCGATTTGGCAAAAGGTGCAAGACCTGTGAAGCACTGCAATCTTAACTTAACTGAG 1564
Db 80815 GCGGCGATTTGGCAAAAGGTGCAAGACCTGTGAAGCACTGCAATCTTAACTTAACTGAG 80874
Qy 1565 AAGTCATCGTCCCTGTCGACCAACGCCAGATGAATTAAGAAAGTATGATGCTGAGATTA 1624
Db 80875 AATTCATTTGCTCTGTCGACCAACGCCAGTGAATTAAGAAAGTATGATGCTGAGATTA 80934
Qy 1625 TCGGCGACTTCTTTGTCAGCAGACTGACAGCGCAAGATCAAGGAAATGTATACAGAG 1684
Db 80935 TCGGCGACTTCTTTGTCAGCAGACTGACAGCGCAAGATATAGGGAATCTGTACAGAG 80994
Qy 1685 TGAAGCAGAGAGAGAAATATACCTCAGAGAGTGCCTCAGAGCGACAGATGAGG- C 1743
Db 80995 TGAAGCAGAGAGAGAAATATACCTCAGAGAGTGCCTCAGAGCGACAGATGAGGCG 81054
Qy 1744 TCCCAAGGCAACAGCAAAACACGATGAAATGTAGCCCTTCCAACTCTGACAGATTA 1803
Db 81055 TCCCAAGGCAACAGCAAAACACAGATGAAATGTAGCCCTTCCAACTCTGACAGATTA 81114
Qy 1804 GACCAAAAGCAGCCAGCCAGATTCGGGAGCAAAACCAAGACCATCTGAGGATGAGAGTC 1863
Db 81115 GACCAAAAGCAGCCAGCCAGATTCGGGAGCAAAACCAAGACCATCTGAGGATGAGAGTC 81174
Qy 1864 TCGGAGCGCGCAGAGACTGCGCAGAGCCCTGAGAACCCCGAGGGCCGAGAGGGCGG 1923
Db 81175 TCGGAGCGCGCAGAGACTGCGCAGAGCCCTGAGAACCCCGAGGGCCGAGAGGGCGG- G 81233
Qy 1924 GGGAGGTGAGCGAGTTTGCAGAACACGAGGCCCGGCTCCCGCCCGCAGAGGCTTC 1983
Db 81234 GGGAGGTGAGCGAGTTTGCAGAACACGAGGCCCGGCTCCCGCCCGCAGAGGCTTC 81293
Qy 1984 TGCAGGCTTCAAGCATTCACATTCACATTCGAGTCTCTGAACTCCACAGAGCG 2043
Db 81294 TGCAGGCTTCAAGCATTCACATTCACATTCGAGTCTCTGAACTCCACAGAGTCG 81352
Qy 2044 TATCCCTTTAGTTGAACATGATAGTGAACGTGTTCAGAACCAAGCAAAATGCAACCC 2103
Db 81353 TGTACTTTTGTGTGAATGATAGTGAACGTGTTCAGAACCAAGCAAAATGCAACCC 81412
Qy 2104 CTTTCTTGTGCAAAATGCTCTGTATCATGTGTGTATATTAAGAGGAAAGTGTAT 2163
Db 81413 CTTTCTTGTGCAAAATGCTCTGTATCATGTGTGTATATTAAGAGGAAAGTGTAT 81468
Qy 2164 AGATATGTGGCTGTGGGTTTACACAGGAGCTGACAGCGTATATATTTAGAAATAAT 2223
Db 81469 AGATATGTGGCTGTGGGTTTACACAGGAGCTGACAGCGTATATATTTAGAAATAATC 81528
Qy 2224 ATATCAATTAATCACTAATCTAATCTTATATCAATTAATTAATTTTCTTTTAT 2283
Db 81529 ATATCAATTAATCACTAATCTAATCTTATATCAATTAATTAATTTTCTTTTAT 81588
Qy 2284 AAGAGAAAGAGGCTTTTCTAAGCTTTAAGAAATAAGCTTTGGAGGCTCAGCGTGT 2343
Db 81589 AAGAGAAAGAGGCTTTTCTAAGCTTTAAGAAATGCT---TTCTTGGAGGCTCAGCGTGT 81644
Qy 2344 AGAGAGAGCTTTGAGGACACCGCACAAATCAACCAAGAGGAAATCTGTGAGAG 2403
Db 81645 AGAGAGAGCTTTGAGGAGTACCTACCAAAATTTCAACCAAGAGGAAATCTGTGAGAG 81704
Qy 2404 ACACTCAGGCAATTTCTGATCACTGTGTATGTCAACAGAGGATACGCTCTCTTGA 2463
Db 81705 ACACTCAGGCAATTTGATGATCTGTGTATATCAACAGAGGATACCATCTCTTGA 81764
Qy 2464 AAGAGAACTGTCTCACTCTCTATGCTGTGATGCTATACCAACATTTCTCTTGTTC 2523
Db 81765 AAGAGAACTGTCTCACTCTCTATGCTGTGATGCTATATCAATTTCTCTTGTTC 81824
Qy 2524 ACAGGTTTAACTGTGTTTGTGATATGCTATATATCTC--TGCTCTCTGTCTT 2581
Db 81825 ATGGGTTTAACTGTGTTTGTGATATGCTATATATCTC--TGCTCTCTGTCTT 81884
Qy 2582 ATCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2641
Db 81885 ATCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 81944
Qy 2642 CTTTCATCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2693
Db 81945 CTTTCATCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 82004
Qy 2694 GCTCTGATTCATCATCAACAAAGGAAACAAAGGCAACAAACGAGCTCTCACT 2753
Db 82005 GCTTGTGATTCCTCAACAAAGGAAACAAAGGCAACAAACGAGCTCTCACT 82064
```


repeat_region	/rpt_family="TLMnC"	/evidence=not_experimental
repeat_region	12790.13085	/rpt_family="AluUb"
repeat_region	complement(13086..14349)	/evidence=not_experimental
repeat_region	/rpt_family="TLMnC"	14367..14502
repeat_region	/evidence=not_experimental	/rpt_family="TLMb"
repeat_region	14724..14905	/evidence=not_experimental
repeat_region	/rpt_family="TLMc"	15003..15180
repeat_region	/evidence=not_experimental	/rpt_family="TLMd"
repeat_region	15217..15433	/evidence=not_experimental
repeat_region	/rpt_family="TLMb3"	16753..16837
repeat_region	/evidence=not_experimental	/rpt_family="AT_rich"
repeat_region	17096..17145	/evidence=not_experimental
repeat_region	/rpt_family="Tigger1"	17146..17170
repeat_region	/evidence=not_experimental	/rpt_family="(Tg).n"
repeat_region	17171..18147	/evidence=not_experimental
repeat_region	/rpt_family="Tigger1"	18148..18427
repeat_region	/evidence=not_experimental	/rpt_family="AluSg"
repeat_region	18428..19574	/evidence=not_experimental
repeat_region	/rpt_family="Tigger1"	complement(19575..19852)
repeat_region	/evidence=not_experimental	/rpt_family="AluSx"
repeat_region	19853..20035	/evidence=not_experimental
repeat_region	/rpt_family="Tigger1"	20272..20565
repeat_region	/evidence=not_experimental	/rpt_family="AluSg"
repeat_region	complement(20853..20937)	/evidence=not_experimental
repeat_region	/rpt_family="TTLH"	complement(21626..21927)
repeat_region	/evidence=not_experimental	/rpt_family="AluSx"
repeat_region	complement(22312..24467)	/evidence=not_experimental
repeat_region	/rpt_family="pIR5"	complement(24463..24524)
repeat_region	/evidence=not_experimental	/rpt_family="HERV9"
repeat_region	complement(24523..24823)	/evidence=not_experimental
repeat_region	/rpt_family="HATLeguin"	complement(24954..25233)
repeat_region	/evidence=not_experimental	/rpt_family="TLMa8"
repeat_region	complement(25679..25860)	/evidence=not_experimental
repeat_region	/rpt_family="MER58D"	complement(25861..26177)
repeat_region	/evidence=not_experimental	/rpt_family="AluSg"
repeat_region	complement(26178..26364)	/evidence=not_experimental
repeat_region	/rpt_family="MER58D"	

```
repeat_region 26399. .26798  
/evidence=not experimental  
/rpt_family="T1"  
repeat_region 27355. .27395  
/evidence=not experimental  
/rpt_family="AT_rich"  
repeat_region 27958. .27979  
/evidence=not experimental  
/rpt_family="AT_rich"  
repeat_region 29533. .29562  
/evidence=not experimental  
/rpt_family="AT_rich"  
repeat_region 30365. .30397  
/evidence=not experimental  
/rpt_family="AT_rich"  
repeat_region 30655. .30712  
/evidence=not experimental  
/rpt_family="AT_rich"  
repeat_region 30881. .30935  
/evidence=not experimental  
/rpt_family="AT_rich"
```

Query Match	70.8%;	Score 2325.4;	DB 9;	Length 113201;
Best Local Similarity	89.3%;	Pred. No. 0;		
Matches 2725; Conservative	0;	Mismatches 212;	Indels 113;	Gaps 16;

QY	247	GGTAAAGTGAATTCATGATGGGAAATTCATGAAAGTGTGATACCTGAGTCTTAAAGACCTA	306
Db	94347	GGTAAAGTGAATTCGATGGGAAATTCATGGAAGTGTGATCTCGGTCTTAAAAACCTA	94288
QY	307	AGGAGAGAGAAATTCAGTTGCGAAACATCCCTCTCACTCGCAGTGGAGAGTGTGAT	366
Db	94287	AGGAAACAGAAATTCAGTTGAAACATACCTCTCACTCGCAGTGGGCGATGTGTGAT	94228
QY	367	GGACTTTTGGCTCAATATGGAACAAGTGAAGATGTGAAACAAGTCACACAGACACGA	426
Db	94227	GCACCTTTGGCTCAACAACGGGATATGTGAGAGATGTGAAACAAGTCACACAGACAGAA	94168
QY	427	ACCGCCGTTGTCAACGTCACATATCCACAAGAGAAAGCAAAATATGCCATGAGAG	486
Db	94167	ACGCTGTTGTCAATCTCATATGCAACAAAGGAAAGACAAATATGCCATGAGAGAG	94109
QY	487	CTAAGCGGGCATCAGTTTGAAGACTACTCTTCAAAGATTTCTACATCCGAGATGAAG	546
Db	94108	CTAAGCGGGCATCAGTTTGAAGACTATCTTCAAAGATTTCTACATCCTGATGAAGAG	94049
QY	547	GTTGAGCT - CCCCTTCCGCCCCCTCAGCGAGCCAGGTGGGACCATCTCCCGGAGCA	605
Db	94048	GTTGAGCTCCCCCTTTACCCCCCAGGAGCCAGGATGGGACCACTCTT - CCGGAGCA	93990
QY	606	AGGCGACGCCCCCTGGGGGCACTTCTCAGCGCAGACAGATTGATTTCCGCTCCGATCCT	665
Db	93989	AGGTCACG - CCGGGGGGCTCTTCTCAGGCCGAGACAGACTGATTTCCGCTCCGATCCA	93931
QY	666	GGTCCCAACCCAGTTTGTGTGTGTCATATCGGAAAGAGAGGCTTGACCATTAAGAAAT	725
Db	93930	-----ACCCAGTTTGTGTGTGCCATGGCTCTGAAAGGAGGCTTGACCATATGTGAAT	93878
QY	726	CACCTAACAGACCCAGTCCCGGGTAAATATTCATATGAAAGAAAGAACTCGAGCTCAGA	785
Db	93877	CACCTAGACAAACCCAGTCCCGGGTAAACATCATTAAGAAAGAACTCGAGGCTCAGA	93818
QY	786	GAAGCCTGTCAACATCCATGCAACCCCAAGAGGGAATTCTGAAAGCATGCCCATATCT	845
Db	93817	GAAGCCTGTCAACATCCATGCCACCCCAAGAGGGGAATTCTGAAAGCATGCCACATCT	93758
QY	846	TGAATATATCAAGAAAGAGCAATGAGCAAACTAGACGCAAGAGATTCTCTGAAAT	905
Db	93757	TGAATATATCAAGAAAGAGCAATGAGCAAACTAGACCAAACTAGCGCAAGAGATTCTCTGAAAT	93698
QY	906	CTTGCGACACAAATGCTGTGTTGGAAGACTGATTGAAAGAAAGGACAGAAATTTGAAGA	965
Db	93697	CGTGGCCACATGCTGTTGTTGGAAGACTGATTGGAAGAAAGAGGACAGAAATGTGAAGA	93638

QY	966	AATTGAACATGAAAACAGGGACCCAGATTAACAATCTCATCTTTGACAGATTGGACATTA	1025
Db	93637	AAATGAACATGAAAACAGGGACCAAGATTAACAATCTCATCTTTGACAGATTGGACATTA	93578
QY	1026	CAATCCGGAAAAGAACCAATCACTGTGGAAGGACAGATTAGGCTGTGCCAGTCTGAAGAT	1085
Db	93577	CAATCCGGAAAAGAACCAATCACTGTGGAAGGACAGATTAGGCTTATGCGACGTGAGAT	93518
QY	1086	AGAGATTATGAAGAAGCTGCGTAGAGCCCTTGGAAAATGATATGCGTGGCTTTAAACCCA	1145
Db	93517	GGAGATTATGAAGAAGCTGCGTAGAGCCCTTGGAAAATGATATGCGTGGCTTTAAATACCA	93458
QY	1146	CTCCGGATAC-TTCTCCAGCCTGTATACCCCATCAACAGTTTG3CCGCTTCCCGCATATC	1204
Db	93457	CTCCAGATATCTTCTCCAGCCTGTATACCCCATCAACAGTTTG3GCCGCTTCCCGCATAGC	93398
QY	1205	ACTCTTATCCAGACAGGAAGTTTGATCTCTTCATCCCAACCCAGGCTGTG3CGCCA	1264
Db	93397	ATTCTTATCAAGACAGGAAGTTTGATCTCTTCATCCCAACCCAGGCTGTGGGTACCA	93338
QY	1265	TCATCGGAAAAGAGGGGGACAACAATCAACGCTGG3AGATTTGGCCGAGACCTTATCA	1324
Db	93337	TCATCGGAAAAGAGGGGGACAACAATCAACGCTGG3AGACTGGCCAGACCTTCA---	93281
QY	1325	AGATTGCCCTGCGGAAGGCCACGTCAGCGAAGGATGGTCATCATCACCG3GCAC	1384
Db	93280	-----	93281
QY	1385	CGGAAGCCAGTTCAAGGCCCAAGGACCGATCTTTGG3AAACTGAAAAGGAAAACTTCT	1444
Db	93280	-----TCAAGGCCCAAGGACCGATCTTTGG3AAACTGAAAAGGAAAACTTCT	93233
QY	1445	TTAACCCTCAAGAGAAGTGAAGCTGGAACGGCATACAGAGTSCCTCTTCCACAGTGG	1504
Db	93232	TTAACCCTCAAGAGAAGTGAAGCTGGAACCCCATATCAAGTCTCTTCCACAGTGG	93173
QY	1505	GCCGGGTGATTGGCAAAAGTGG3ACAGACCTGTAAACGATACGACAACTTAACTGAAGTGG	1564
Db	93172	GCCGGGTGATTGGCAAAAGTGG3ACAGACTGTGAATGAATCTGTAACTTAACTTATGTGG	93113
QY	1565	AAGTATCTGCTGCTGCTGACCAAAACGCAGATGAAAAATGAGAAAGTATCGTCAGAATTA	1624
Db	93112	AAATCATTTGTGCTGCTGACCAAAACCCCATGAAAAATGAGAAAGTATCATCAGAATTA	93053
QY	1625	TGCGGACACTTTTGTGTAGCCACATGACACAGGCGAAGATCAGGAAAATTGTACAAAGG	1684
Db	93052	TGCGGACACTTTTGTGTAGCCACATGACACAGGCGCAAGTATCAGGAAACTGTACAAAGG	92993
QY	1685	TGAAGCAGCAGAGACGAATAATCCCTCAGG3AGTGGCTCTACAGCGCAGCAAGTGAAG-C	1743
Db	92992	TGAAGCAGCAGAGAGTGAATAATCCCTCAGG3AGTGGCTCTACAGCGCAGCAAGTGAAGCC	92933
QY	1744	TCCCAAGGCAACCAAGCAAAACAAACGATGATGTAGCCCTTCCAAACACTGACAGATGA	1803
Db	92932	TCCCAAGGCAATCAGCAAAACAAACGATGATGAAGCCCTTCCAAACACTGACAAATGA	92873
QY	1804	GACCAAAACGACACGACCCAGATGTGG3AGCAAAACAAGACATCTGAGATGAAGTGC	1863
Db	92872	GACCAAAACAACCCACGACCCAGATGTGG3AGCAAAACAAGACATCTGAGATGAAGTGC	92813
QY	1864	TGCG3AGGCG3CAGG3GACTGTG3CG3AGG3CCCTGTAG3AAACCC3AG3GG3CG3AG3G3G	1923
Db	92812	TGCG3AGGCG3CAGG3GACTGTG3CG3AGG3CCCTGTAG3AAATCC3AG3GG3CA3AG3G3G-G	92754
QY	1924	GG3AAAGTCAAGCTTGTGCAAGAACCAACGACCCCGCTTCG3CCCC3AG3GCTTC	1983
Db	92753	GG3AAAGTCAAGCTTGTGCAAGAACCTTCG3GGCCCTGTCTTAOCTTCCAGACTTC	92694
QY	1984	TGC3AGGCTTCAAGCCATCCACTTCAACATCACTCG3ATCTCTCTG3AATCTCC3AG3CGC	2043
Db	92693	TGCA-GCTTCAAGCAATTCATTCACATCACTCAATCACTCAATCTCTCTTAATCTCC3CAATGC	92635

QY	2044	TAATCCCTTTTAGTTGAACCTAAATAGTGAACGTGTTCAAGGCCAAGCAAAATGCACACC	2103
Db	92634	TCGTACCTTTTGTTGTAAGTAACATAGGGAACGTGTTCAAGCCATGCAAAATGCACACC	92576
QY	2104	CTTTTCTGTGCAAAATGCTCTCTGTACATGTGTGTACATATTGAAAGGGAAGATGTTA	2163
Db	92574	CTTTTGTGTGCAAA----TCTCTGTACATGTGTGTACATATTGAAAGGGAAGATGTTA	92511
QY	2164	AGATATGTGCGCTGTGCGGTACACAGGGTGCTGTACAGGGTAATATATTATTAGAAATAT	2223
Db	92518	AGATACGTGCGCTGTGCGGTATCGAAGGTGCTGTACAGGGTAATATATTATTAGAAATAC	92455
QY	2224	ATATCAAAATTAACCTCAACTCAATTTTAACTAAATTAATTTTATTTTCTTTTCTTTTA	2283
Db	92458	ATATCAAAATTAACCTCAACTCAATTTTAACTAACTAATTAATTTTCTTTTCTTTTAA	92391
QY	2284	AAGAGAAAGCAGCGCTTTCTTGACTTTAAAGATTAAGTCTTTGGAGGTCTACCGGTGT	2343
Db	92398	AAGAGAAAGCAGCGCTTTCTTGACTTTAAAGTCTTTTAAAGTCT----TCTTTGGAGGTCTACCGGTGT	92344
QY	2344	AGAGAGAGCTTTTAGGCGCACCGCACAAAATTAACCAAGGGAAATCTGTGTGGAAAGS	2403
Db	92342	AGAGAGAGCTTTTAGGCGGTGACCTTACAAAATTAACCAAGGGAAATCTGTGTGGAAAGS	92281
QY	2404	ACACTCAGCGAGTTTGATCACTGTGTATGTCAACAGAGGATACCGTCTCTGTGA	2463
Db	92282	ACACTCAGCGAATTTAGATCACTGTGTATGTATCAACAGAGGATACCATCTCTGTGA	92222
QY	2464	AGAGGAAACTGTGCATCTCTCATGTCTGTGTACATACCCATTTCTCTTGTCTTC	2523
Db	92222	AAAGGAAACTGTGCATCTCTCATGTCTGTGTATACATCAATTTCTCTTGTCTTC	92163
QY	2524	ACAGGTTTAAACGTGTTTWTGATACGTGTATTAATATCTG--TGTCTGTCTGTGT	2581
Db	92162	ATGTGTTTAAACGTGTTTWTGATACGTGTATTAATATCTGATCTGTCTGTGTGT	92103
QY	2582	ATCTCTCCCTCCCTCCCTCCCTCTCTCTCATCTCCATCTTTTGTGAATTTCTCTCATC	2641
Db	92102	ATCTCTCACCTCCCTCCCTCCCTCCCTCATCTCCATCTTTTGTGAATTTCTCTCATC	92043
QY	2642	CCTTCATCTCAATCCGTATCTTAGCGA-----CCCCCCCCCCCCAGGCAAGCAGT	2693
Db	92042	CCTTCATCTCAATCCGTATCTTAGCGAACACACACCCCTGCGCCCCGCAAGGAAACAGT	91983
QY	2694	GCTCTGAGTATTCACATCACACAAAAGGAACAAAAGGAAACACACAAACAGCTCAACT	2753
Db	91982	GCTTTGAGTATCCCAACACACAAAAGGAACAAAAGGAAACACACACAGTCAAGCTCAACT	91923
QY	2754	TACACTTGTTACTCAAAAAGAACAAAGTCAATGTATCTGTCTCAAGCTTTTGGAAAGS	2813
Db	91922	TAGCTTGTTATCTCAAAAAGAACAAAGTCAATGTATCTGTCTTGACATTTTGGAAAGS	91863
QY	2814	GAAAACGGAACCCACCAACCAACCAATCAACCAACCAAGAAAAAATTTCACAATGAA	2873
Db	91862	GAAAACGGAACCCACCAACCAACCAATCAACCAACCAAGAAAAAATTTCACAATGAA	91803
QY	2874	AGATGTATTTTGTCTTTTGTGATTTGTGTATTAAGCATATATTACAGAAATGAT	2933
Db	91802	AGATGTATTTTGTCTTTTGTGATTTGTGTATTAAGCATATATTACAGAAATGAT	91743
QY	2934	TCCTTTCTTTAAAAAAAATGTGAGGAAAGTGAATTTTACCAAGGTTGTGGCCCA	2993
Db	91742	TCCTTTCTTTTAAAAAATGTGAGGAAAGTGAATTTTATCAAGGTTGTGGCCCA	91683
QY	2994	GGGCGTTAAATTCACAGATTTTATTAACAGAAAAACACACAGAAAGACTCAAGGT	3053
Db	91682	GGGCGTTAAATTTTACAGATTTTATTAATGAGAAAAACACACAGAAAGACTCAAGGT	91623
QY	3054	G-TTTTAACTGAGACCTTGCTGTGTGTCTCCCTTGAAGTTTTGTGTAAGCTCATAGT	3112
Db	91622	GTTTTTGTCTGACACCTTGCTGTGTGTCTCTTAGAGATTTTGTGTAAGCTCATAGT	91563
QY	3113	TGAGGCAATTTTATTTTATTTTAAATTAATATAGTTGGAAGAAAAAATATAGATATCAATG	3172

Db 91562 TGGAGCA-----TTATTTCTTAATAAATGAGTGGAAAAAATAGATATACAA 91507
QY 3173 CCAGCTTGAGAGAGGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATGTTCTCCGCTA 3232
Db 91506 CCAGCTTGAGAGAGGTGAGAGTCCAAATGTGTGCAACAGCTGTTCTGAATGTTCTCCGCTA 91447
QY 3233 GCCAAGAACCMATATGCGCTTCTTTTGGACCAACCTTGAAATGTTTMT 3282
Db 91446 GCCAAGAACCTATATGCGCTTCTTTTGAACAACCTTGAAACATTAT 91397
RESULT 14
BC023758 3557 bp mRNA linear ROD 16-APR-2003
LOCUS Mus musculus RIKEN cDNA C330012H03 gene, mRNA (cDNA clone
IMAGE:5354659), partial cds.
ACCESSION BC023758
VERSION BC023758.1 GI:23958572
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3557)
AUTHORS Strausberg,R.L., Peltingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schectel,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carinci,F., Frange,C., Raha,S.S., Loggellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McKean,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,B.J., Lu,X., Gibbs,R.A.,
Fahy,U., Helton,E., Kettelman,M., Madan,A.C., Shcherchenko,Y.,
Sanchez,A., Whiting,M., Madan,R., Touchman,D.W., Green,E.D.,
Boutard,G.G., Blakesley,R.W., Touchman,D.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.B.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
TITLE
JOURNAL
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 3557)
AUTHORS Strausberg,R.
DIRECT SUBMISSION
Submitted (05-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Contact: nisc.mgc@hgti.nih.gov
Web site: http://www.nisc.nih.gov/
REMARK
COMMENT
Ahther,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Boufard,G.G., Brein,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-U., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantirpop,S., Thomas,P.J., Touchman,J.W.,
Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,

Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAC Plate: 55 Row: 0 Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Genomescan gene prediction.
location/Qualifiers
1. 3557
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5354659"
/tissue_type="Mammary tumor, C3(1)-Tag model, infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="NCI_COAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
ORIGIN
Query Match 58.6%; Score 1924.2; DB 10; Length 3557;
Best Local Similarity 79.7%; Pred. No. 0; Mismatches 459; Indels 229; Gaps 28;
Matches 2708; Conservative 0;
58 GGAAGAGCGAGATGATGAACAAGCTTTACATCGGAAACCTGAGCCCGCGCTACCGCC 117
Db 23 GGAAGAGCGAGATGATGAACAAGCTTTACATCGGAAACCTGAGCCCGCGCTACCGCC 82
QY 118 GAGCACTTCGCGAGCTCTTTGGGAGCAGAAAGCTCCCTGCGGAGCAGGTCCTGCTG 177
Db 83 GAGCACTTCGCGAGCTCTTTGGGAGCAGAAAGCTCCCTGCGGAGCAGGTCCTACCTC 142
QY 178 AAGTCGCGTACGCTCTGTGTGACTACCCGACCAAGATCGGACATCGCCCATCGAG 237
Db 143 AAGTCGCGTACGCTCTGTGTGACTACCCGACCAAGATCGGACATCGGACATCGCCCATCGAG 202
QY 238 ACCCTCTCGGTTAAAGTGAATGCAATGGAATCGGAAATCATGGAAGTTGATTACTCACTCTCT 287
Db 203 ACCCTCTCGGTTAAAGTGAATGCAATGGAATCGGAAATCATGGAAGTTGATTACTCACTCTCT 262
QY 298 AAAAGCTAAGAGCAGGAAATTCAGATTCGAAACATCCCTCCACCTGCGAGTGGAG 357
Db 263 AAAAGCTAAGAGCAGGAAATTCAGATTCGAAATCCCGCTCACTGCGAGTGGAG 322
QY 358 GTGTGGATGGAATTTGGCTCAATATGAGACAGTGAATGTGGAACAAGTCAACACA 417
Db 323 GTGTGGATGGAATTTGGCTCAATATGAGACAGTGAATGTGGAACAAGTCAACACA 382
QY 418 GACACAGAAACCGCCGTTGTCAACGTACATATGCAACAGAGAGCAAAATRAGCC 477
Db 383 GATACAGAAACCGCCGTTGTCAACGTACATATGCAACAGAGAGCAAAATRAGCT 442
QY 478 ATGGAAGATTAAGCGGCGATCATGTTGGAACATCACTCTCAAGATTCTTCATCCCG 537
Db 443 ATGGAAGATTAAGCGGCGATCATGTTGGAACATCACTCTTCATCAAGATTCTTCATCCCG 502
QY 538 GATGAAGAGTGAAGTCCCTTCGCCCCCTCAGCCAGCCAGCGAGTGGGAGCACTCTCC 597
Db 503 GATGAAGAGTGAAGTCCCTTCGCCCCCTCAGCCAGCCAGCGAGTGGGAGCACTCTCC 541
QY 598 CGGAGCAAGAGCCAGCCCTCGGGGCACTTCTCAGGCGAGCAAGATTTGATTTCCGCTG 657
Db 542 CGGAGCAAGAGCCAGCCCTCGGGGCACTTCTCAGGCGAGCAAGATTTGATTTCCGCTG 601
QY 658 CGGATCTGATCCCAACCCAGTTGTTGGTCATCATGGAAGAGGAGGCTTGACCAT 717
Db 602 CGGATCTGATCCCAACCCAGTTGTTGGTCATCATGGAAGAGGAGGCTTGACCAT 661
QY 718 AAGAACTACTAGACAGACCCAGTCCCGGATGATATCAATGAAAGAGAACTCTGGA 777
Db 662 AAGAACTACTAGACAGACCCAGTCCCGGATGATATCAATGAAAGAGAACTCTGAG 721

```
QY 778 GCTGAGAGAGAGCTGTACCATTCATGACCCCGAGAGGAGCTTTGAAAGCATGCGCC 837
Db 722 GCTGAGAGAGAGCTGTACCATTCATGACCCCGAGAGGAGCATTTGAAAGCATGCGCC 781
QY 838 ATGATTTCTTGAATCATGCAAGAAAGGAGGAGTGAACCAATAGCCGAGAGATTCT 897
Db 782 ATGATTTCTTGAATCATGCAAGAAAGGAGGAGTGAACCAATAGCCGAGAGATTCT 841
QY 898 CTGAAGATCTTGGGACACAAATGGCTTGGTGGGAGACATGATTGGAGAAAGAGCGAAT 957
Db 842 CTGAAGATCTTGGGACACAAATGGCTTGGTGGGAGACATGATTGGAGAAAGAGCGAAT 901
QY 958 TTGAAGAAATTTGAACATGAAACAGAGGACCAAGATTAACATCTCATCTTTTGAGAGATTG 1017
Db 902 CTGAAGAAATTTGAACATGAAACAGAGGACCAAGATTAACATCTCATCTTTTGAGAGATTG 961
QY 1018 AGCATATTCAAACCCGAGAAAGAACATCATCTGTGAAGGGGACAGTTGAGGCTGTGCCAGT 1077
Db 962 AGCATATTCAAACCCGAGAGAAACATCATCTGTGAAGGGGACAGTTGAGGCTGTGCCAGT 1021
QY 1078 GCTGAGATGAGATTATGAGAGAGTGGGTGAGGCTTTGAAATGATATGCTGGCTGT 1137
Db 1022 GCTGAGATGAGATTATGAGAGAGTGGGTGAGGCTTTGAGAAAGATGCTGGCTGT 1081
QY 1138 AAC----- 1140
Db 1082 AACCAACAGCCATCTGATCCAGGGGTAAACCTCACTGACATTTGGCATTTTCCACT 1141
QY 1141 ----- 1140
Db 1142 GGACTGTCTGTCTCTCCACAGAGAGGCCCGTGGAGTTCCCGCCAGTCTCCCTAT 1201
QY 1141 -----ACCACTCCGGAATCTTCTCCAGGCTGTACCCCATACCAATTTGGC 1188
Db 1202 CACCCCTTTGGTACCACTCCGAACTTCTCCAGTCTGTACCCCTCATCAATTTTGGC 1261
QY 1189 CCGTTCCCGCATCATCACTCTTATCCAGAGAGAGATTGGATCTCTTCAATCCCAACC 1248
Db 1262 CCATTCACATCATCACTCTTATCCAGAGAGAGATTGGATCTCTTCAATCCCAACC 1321
QY 1249 CAGGCTGTGGGCGCATCATCGGGAGAGAGGGGACACATCAAAACAGCTGGCGAGATTC 1308
Db 1322 CAGGCTGTGGGCGCATCATCGGGAGAGAGGGGACACATCAAAACAGCTGGCGATTC 1381
QY 1309 GCGGAGGCTCTTATCAAGATTGGCCCTCGGAGAGGCCCAAGCTGACGAAAGGATGTC 1368
Db 1382 GCTGTGTCTCTCAATCAAGATTGGCTCGAGAGAGGTCAGATGTCAATGAGAGATGTC 1441
QY 1369 ATCATCACCGGCGCACCGAGAGCCAGTTCAAGGCCAGGGAAGGATCTTTGGGAAACTG 1428
Db 1442 ATCATCATGCTCTCTCTGAAGCCAGTTTAAGGCTCAGGGAAGGATCTTTGGGAAACTG 1501
QY 1429 AAAAGAGAAACTTTCTTAAACCCCAAGAAAGTGAAGCTGGAAGGCAATATCAGAGTG 1488
Db 1502 AAGAGAGAAACTTTCTTAAATCCCAAGAAAGTGAAGCTGGAAGGCCCAATCCAGATC 1561
QY 1489 CCGCTTTCCACAGCTGGCGGGGTGATTGGGAAAGGTGGCAAGACCTGTGAACGATCTGAG 1548
Db 1562 CCGCTGTCCAGCGCTGGCGGGGTGATTGGGCAAGGGCGGGAAAAACCTGTGAACGATCTGAG 1621
QY 1549 AACTTAACAGTGCAGAGTCAATGCTGCTGTGACCAACGCGCAGATGAAATGAGGAA 1608
Db 1622 AACTTAACAGTGCAGAGTCAATGCTGCTGTGACCAACGCGCAGATGAGGAA 1681
QY 1609 GTGATGTGAGATTATCGGGCACTTTCTTGTAGCCAGACTGACAGCGCAAGATCAGG 1668
Db 1682 GTGATGTGAGATTATCGGGCACTTTTGTAGCCAGACTGACCAAGCAAGATCAGG 1741
QY 1669 GAAATTTGTCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1728
Db 1742 GAAATTTGTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1801

QY 1729 CGCAGCAAGTGAAGGCTCCACAGGACACAGCAAAACACGATGATGATGAGCTTCCAA 1788
Db 1802 CGCAGCAAGTGAAGGCTCCACAG--CACACAGCAAAACGATGATGATGAGCTTCCAA 1859
QY 1789 CACCTGACAGATGAGACCAACAGCAGCAGCAGATGAGGAGGAGCAAAAGACAT-C 1847
Db 1860 CACCTGACAG-ATGAGACCAAC--AGCAGCAGATGAGAGCAAAACCAAGACATCC 1915
QY 1848 TGAGGATGAGAAAGTCTGCGAGGCGGCGAGGAGACTGCGAGAGGCGCTGAGAACCCAG 1907
Db 1916 CGAGAGTGGCGCAGATCTGCAAGACAGCCAGGCGCTGC-----AGACCTTACATCTCG 1970
QY 1908 GGGCGGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1965
Db 1971 GGAATCGAGAGGAGCAAGGAG--GCCAGGTTGTCCAGAAACACCGCTTGGCGCTGCGC 2027
QY 1866 CCGCGCCCGCCAGGCTTCTGAGGCTT--CAGCCATCCACTTACATCCACTGCGATCT 2023
Db 2028 CCAGCTTCCCTGGCTTCTGAGGCTTACAGGCAATCCACTGCGCTCATCACTCAGAT-T 2086
QY 2024 CTCTGAACTCCACAGAGCTATCCCTTTTATGTTGAATMAATAGTGAACGTTGCA 2083
Db 2087 CTCTCAGTTCCAGAGAGCTATCCCTTGGTTGAATMAATAGTGAACGTTGCA 2146
QY 2084 AGCCAAGCAAAATGCACACCTTTTC-----TGTGCAAAATGCTCTGTATGATGTGT 2139
Db 2147 AGCAAGCAAAATGCTAGCGCTTCTTGTGGAAGAGTGTCTGTATCATATGATGT 2206
QY 2140 ACATATTGAAGAGGAA-GATGTTAATATGTTGGCTGTGAGGTTTACAGAGGCGCTGC 2198
Db 2207 ACATATTGAAGAGGAGAGATGTTAAGAAATGTGCTGTGAGGTTTACAGAGGCGCTGC 2266
QY 2199 AGCGTATATATTTTGAAGAAATATATATCAATTAATCAATCAATTTTATATC 2258
Db 2267 AGCGTATATATTTTGAAGAAATATATATCAATTAATCAATCAATTTTATATC 2326
QY 2259 AATATTAATTTTTTTTCTTTTAAAGAGAGAGGCTTTTATGACTTTTAAAGATA 2318
Db 2327 AGTATTAATTTTTTTTCTTTTGAAGAGAGGCTTTTATGACTTTTAAAGATA 2386
QY 2319 AGCTCTTGG-GAGGTCTCAGCGGTGAGAGAGGCTTTGAGGC-CACCGGCAAAAT 2376
Db 2387 AAGATTGGGTGAGGTCTCAGGTATGAAAGGACTTTGAGCCACCAAAAT 2446
QY 2377 CACCCAGAGGAAATCTGTGAGAGAGCACTCAGCAGTTCTGATCACTGTATG 2436
Db 2447 CACTTGGAGGAAATCTGTG-AAGAGCACTTAATGCAATGAGGTGAGGTGAGGTG 2505
QY 2437 TCAACGAGAGGATPACGCTCTCTTGAAGAGAACTCTGCTCATGCTGTCTA 2496
Db 2506 TCAACGAGAGGATPACG--TCCCTGAGAGGAGGCGCTGAGCGCTGTCTCTC 2563
QY 2497 GCTCATACCCCATTTCTCTT-----TGCTCAAGGTTTAA 2534
Db 2564 CCACACACCCATGCTCAGTTTTCAGAGTTTAACTGATGATTTTGGTTGGTTGT 2623
QY 2535 ACTGTTTTTTGATACCTGATATATATCTGTCTCTCTCTGTTATCTCTCCCTCC 2594
Db 2624 TTTGTTTTTTGATACCTGATATATATGTTCTGTCTCTC-----CTTATCTCCCTCT 2677
QY 2595 CTCCCTCCCTCTCTCTCTCATCTCAATCTTTTGAATTTCTCATCTCTCATCTCAAT 2654
Db 2678 AACTTCTCTGTCTCTCTTAAACCAAAATCTTTTGAAGTCCCTGTGTTCTTCTCTC 2737
QY 2655 CCGGTATCTACGACG---CCCGCCCGCCAGGCAAGAGCTGTGAGTATCAGATCA 2711
Db 2738 CCGGTATCTATGACCCCTCTCCCTCCCGCCAGGCAAGAGCTGTGAGTATCAGATCA 2797
QY 2712 CACAAAGGAGCAAAAGGAGAAACACAAACGAGCTCACTTACCTTGGTTACTCAAA 2771
Db 2798 CACAAAGGAGCAAAAGGAGAAACACACAGAGCTCAGCTTACCTGGTTACTCAAG 2857
QY 2772 AGAAAGAGTCAATGTATCTTGTCTACGCTTTTGAAGAGGAAACAGAACCAACCA 2831
```

Db 2858 CGAACAAGAGTCGGCGGACTTGTCTTACG-ATTGGAAGAGGAAAGCAGATCCGACGA 2916
QY 2832 AACCAACCAATCAACCAACCAACCAACCAACCAATTCACATGAAGAATGTTGTCTT 2891
Db 2917 AATAACACAGCAGCAACCAACCAAG-AATAATCCCAAGAAAGAAATGGA----CTCTCT 2971
QY 2892 TTGCATTTTGGTGTATTAAGCCATCAATTTACGAAATGATTCCTTCTTTAAAAAAA 2951
Db 2972 CTCATCTTGGTATACAGCCATCAACATTTAGTAAATCTTCTTCTTTAAAAATTT 3031
QY 2992 AAATGTGAGGAAAGTAGAAATTTACAAAGTGTGTGGCCAGGCGCTTAATTCACAGA 3011
Db 3032 AAAGTGCAGGAAATAGCAATTTACAGAGTGTGTGGCCAGGCGCTTAATTCACAGA 3091
QY 3012 TTTTATTAACGAGAAACACACAGA--AGAGTACTCTAGGTG-TTTTACTCTGAC 3067
Db 3092 TTTTATTAATGAGAAACACACAAATTAAGCTACGCTGAGTGTATTTTAACTCTGAC 3151
QY 3068 ACCTTGTCTGTGTCTTCCCTTACAGATTTGTAAAGCTGATGAGTGTGAGCATTTT 3127
Db 3152 ACCTTGTCTGTGTCTTCCCTTACAGATTTGTAAAGCTGATGAGTGTGAGCA-TTTT 3210
QY 3128 TTTTATTAATTAATGATGAGTGTGAGAAATTAATGATATCACTGCGCTGAGAGAG 3187
Db 3211 TTTTATTAATTAATGATGAGTGTGAGAAATTAATGATATCACTGCGCTGAGAGAG 3268
QY 3188 TGACAGTCCAAAGTGTGACACAGCTGTCTGAAATGTCTTCCGCTACCCAGAACCAAT 3247
Db 3269 CGGAGCCGACAGGTGTGCGCAGCTCTCTC---TGTCTTCCGCTACCCAGAGCTAGCT 3323
QY 3248 GGCCTTTCTTTGGACAAACCTTGAAATGTTATTT 3283
Db 3324 GGCCTTTCTTTGGACAAACCTTGAAATGTTATTT 3359

RESULT 15
AF117107 2010 bp mRNA linear PRI 26-JAN-1999
LOCUS Homo sapiens IGF-II mRNA-binding protein 2 (IMP-2) mRNA, complete
DEFINITION cds.
ACCESSION AF117107
VERSION AF117107.1 GI:4191609
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2010)
AUTHORS Nielsen, J., Christiansen, J., Lykke-Andersen, J., Johnsen, A.H.,
Wewer, U.M. and Nielsen, F.C.
A family of insulin-like growth factor II mRNA-binding proteins
represses translation in late development
Mol. Cell. Biol. 19 (2), 1262-1270 (1999)
99108099
JOURNAL MEDLINE
PUBMED 9891060
REFERENCE 2 (bases 1 to 2010)
AUTHORS Nielsen, J., Christiansen, J., Lykke-Andersen, J., Johnsen, A.H.,
Wewer, U.M. and Nielsen, F.C.
Direct Substitution
Submitted (30-DEC-1998) Institute of Molecular Biology, University
of Copenhagen, Soelystgade 83H, Copenhagen DK-1307, Denmark
FEATURES
source
1..2010
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..2010
/gene="IMP-2"
gene
47
/gene="IMP-2"
misc_feature
47
/note="alternative start-codon gives an additional
N-terminal methionine"

CDS
50..1846
/gene="IMP-2"
/note="binds multiple sites in IGF-II mRNA 5'UTR"
/codon_start=1
/product="IGF-II mRNA-binding protein 2"
/protein_id="AB09827.1"
/db_xref="GI:4191610"
/translation="MNLVYGNLSPAVTADLRQLFEDRKPLAGQVLLKSGYAFVY
PDQWALRALETSLGKVELHKLMEVDYSVKLRKRIQIRNIPHLQWVLDGLA
QYGVENVBOVNDPTETAVVNTVATREBAKIAMKLSGHQFENVSEKISYIPDEVS
SPSPQAOARDHDSREOGAPRGSTSOAROIDPLRLILVPOFGAIIIGKGLATKNI
TKQTSRDVIRKNSGAEKRPVTHATPBETSACPMILIEIMKEADETKLAERIPL
KILANGLVGRLLIGKEGNLKKLIEHETGKLTISLDPLSTYNDERITTYGVYACA
SAEIEIMKKLEAFENMDLAVNOQANLIPGLNSALGIFSTGLSVSPFAGKAPPA
APYHFTTHSGYFSLYPHQPGFPFPHHSYRDEIVLPLPTQAVGAIIIGKGAHKK
OLARFAGASIKILADPAGDVSERNVITGPEPAQFKAQRIFGKLEENFENPKBEVK
LEAATRPSSSTAGKGVICKGTAVNELNLNSAEYIVPRDQPDENBEVYVILIGHFA
SQTAKRKIREIVQVKKQEQKRPQGVASQSK"

ORIGIN
Query Match 52.1%; Score 1711.8; DB 9; Length 2010;
Best Local Similarity 92.9%; Pred. No. 0;
Matches 1858; Conservative 0; Mismatches 12; Indels 129; Gaps 1;

QY 35 GCCGGGGAGCCGCGGCTCTTCGGGGAAGAAGCATATGAAACAGCTTTATCATCGGGA 94
Db 12 GCCGTAGGAGACCGCGGCTCAGCGGGAAGAAGCATATGAAACAGCTTTATCATCGGGA 71
QY 95 ACCTAGAGCCCGCGCTCACCGCGACGACCTCCGCGACGCTTTGGGAGACAGAGCTGC 154
Db 72 ACCTAGAGCCCGCGCTCACCGCGACGACCTCCGCGACGCTTTGGGAGACAGAGCTGC 131
QY 155 CCTTGGCGGAGCAGGTCCTGTGTAAGTCCGCTACGCTTGTGTGACTACCCGACAGA 214
Db 132 CCTTGGCGGAGCAGGTCCTGTGTAAGTCCGCTACGCTTGTGTGACTACCCGACAGA 191
QY 215 ACTGAGCCATCCGCGCATCGAGACCTCTCGGGTAAAGTGAATGATGGGAAATGA 274
Db 192 ACTGAGCCATCCGCGCATCGAGACCTCTCGGGTAAAGTGAATGATGGGAAATGA 251
QY 275 TGGAGTGTATTAACACAGCTCTTAAAGCTAAGAGAGCAGAAATTCAGATTCGAAACA 334
Db 252 TGGAGTGTATTAACACAGCTCTTAAAGCTAAGAGAGCAGAAATTCAGATTCGAAACA 311
QY 335 TCCCTCTCACCTGAGAGTGGAGGTGTGATGACCTTTGCTCAATATGGGACAGTGG 394
Db 312 TCCCTCTCACCTGAGAGTGGAGGTGTGATGACCTTTGCTCAATATGGGACAGTGG 371
QY 395 AGAATGTGACACAGTCAACACAGACACGAAACCGCGCTTGCACGTCACATATGGCA 454
Db 372 AGAATGTGACACAGTCAACACAGACACGAAACCGCGCTTGCACGTCACATATGGCA 431
QY 455 CAAGGAAGAAGGAAATATAGCATGAGAGAGTAAAGCGGAGTCAAGTTTGAAGACTACT 514
Db 432 CAAGGAAGAAGGAAATATAGCATGAGAGAGTAAAGCGGAGTCAAGTTTGAAGACTACT 491
QY 515 CCTTCAAGATTTCTACATCCCGATGAAAGAGTGAAGTCCCTTGCCTCCCTCAGCGAG 574
Db 492 CCTTCAAGATTTCTACATCCCGATGAAAGAGTGAAGTCCCTTGCCTCCCTCAGCGAG 551
QY 575 CCGAGCGTGGGACCACTCTTCCCGGAGCAGAGCGACGCGCTTGGGAGCACTTCAGG 634
Db 552 CCGAGCGTGGGACCACTCTTCCCGGAGCAGAGCGACGCGCTTGGGAGCACTTCAGG 611
QY 635 CCAGACAGATGATTTCCGCTCGGATCCTGCTCCCAACCCAGTTTGTGTGACATCA 694
Db 612 CCAGACAGATGATTTCCGCTCGGATCCTGCTCCCAACCCAGTTTGTGTGACATCA 671
QY 695 TCGGAAAGAGGAGCTTGACCATTAAGAACATCACTAAGACAGCCAGTCCCGGTAAGATA 754
Db 672 TCGGAAAGAGGAGCTTGACCATTAAGAACATCACTAAGACAGCCAGTCCCGGTAAGATA 731
QY 755 TCATATGAAAGAGAACTGTGAGCTGACAGAGAGCTGTACATCATCATGCCACCCAG 814

Db 732 TCCATAGAAAAGAGAACTCTGGAGCTGCAGAAAGCCCTGTCACTTCATGCCACCCGAG 791
 QY 815 AGGGGACTTTTGAAGCATTGCCGATATTTTGAATCATGCAAGAAAGGCAAGTTGA 874
 Db 792 AGGGGACTTTTGAAGCATTGCCGATATTTTGAATCATGCAAGAAAGGCAAGTTGA 851
 QY 875 CCAAACTAGCCGAAAGAGATTCCTCTGAAAATCTTGGGCAACAAATGGCTTGGTGGAAAC 934
 Db 852 CCAAACTAGCCGAAAGAGATTCCTCTGAAAATCTTGGGCAACAAATGGCTTGGTGGAAAC 911
 QY 935 TGATTTGAAAAAGAGGCAAGAAATTTGAAGAAATTTGAACATGAAACAGGACCAAGATTA 994
 Db 912 TGATTTGAAAAAGAGGCAAGAAATTTGAAGAAATTTGAACATGAAACAGGACCAAGATTA 971
 QY 995 CAATCTCATCTTTTGCAGATTTTGAAGCATATCAACCCGGAAGAAACCATCTGTGAAG 1054
 Db 972 CAATCTCATCTTTTGCAGATTTTGAAGCATATCAACCCGGAAGAAACCATCTGTGAAG 1031
 QY 1055 GCACAGTTGAGGCTGTGSCCAGTGTGATAGATTAATGAGAAAGCTGGTGAGGCT 1114
 Db 1032 GCACAGTTGAGGCTGTGSCCAGTGTGATAGATTAATGAGAAAGCTGGTGAGGCT 1091
 QY 1115 TTGAAAATGATATGCTGGCTGTAAAC----- 1140
 Db 1092 TTGAAAATGATATGCTGGCTGTAAACCAACAAAGCAATGTGATCCAGGGTTGAACCTCA 1151
 QY 1141 ----- 1140
 Db 1152 GCGCATTGGCATCTTTTCAACAGAGCTGTCCGTATCTCAACAGAGGCCCCGCG 1211
 QY 1141 -----ACCCACTCGGATPACTTCTCCAGCC 1165
 Db 1212 GAGTCCCCCGCGCTGCCCTTACCAACCCCTTCACTACCCACTCCGAAATCTTCTCCAGCC 1271
 QY 1166 TGTACCCCATCAACAGATTTGCCCCGTTCGCGATCACTCTTATCCAGAGCAGAGA 1225
 Db 1272 TGTACCCCATCAACAGATTTGCCCCGTTCGCGATCACTCTTATCCAGAGCAGAGA 1331
 QY 1226 TTGTGAATCTCTTATCTCCAAACCGAGCTGTGGGCGCATCATCGGGAAGAAAGGGGCGAC 1285
 Db 1332 TTGTGAATCTCTTATCTCCAAACCGAGCTGTGGGCGCATCATCGGGAAGAAAGGGGCGAC 1391
 QY 1286 ACATCAAAACAGCTGGCAGATTTGCGCGAGGCTCTATCAAGATTTGCCCTCGGGAAGGCC 1345
 Db 1392 ACATCAAAACAGCTGGCAGATTTGCGCGAGGCTCTTATCAAGATTTGCCCTCGGGAAGGCC 1451
 QY 1346 CAGACGTCAGCGAAAGGATGATCATCAACCGGCGCACCGGAAGCCAGTTCAAGGCC 1405
 Db 1452 CAGACGTCAGCGAAAGGATGATCATCAACCGGCGCACCGGAAGCCAGTTCAAGGCC 1511
 QY 1406 AGGGAACGATCTTTGGGAAACTGAAAGAGGAAAATTCTTTAACCCCAAGAAAGAGTGA 1465
 Db 1512 AGGGAACGATCTTTGGGAAACTGAAAGAGGAAAATTCTTTAACCCCAAGAAAGAGTGA 1571
 QY 1466 AGCTGGAAGGCAATATCAGAGTGCCTTCCACAGCTGGCGGGTATGCGCAAGGTG 1525
 Db 1572 AGCTGGAAGGCAATATCAGAGTGCCTTCCACAGCTGGCGGGTATGCGCAAGGTG 1631
 QY 1526 GCAAGACCGTGAAGCAACTGCAAGACTTAACAGTGCAGAAAGTCAATGTCCTGTAACC 1585
 Db 1632 GCAAGACCGTGAAGCAACTGCAAGACTTAACAGTGCAGAAAGTCAATGTCCTGTAACC 1691
 QY 1586 AAACGCGCAGATGAAGATGAGGAAGTGAATGTCAGAAATTAATCGGCACTTCTTTGCTAGCC 1645
 Db 1692 AAACGCGCAGATGAAGATGAGGAAGTGAATGTCAGAAATTAATCGGCACTTCTTTGCTAGCC 1751
 QY 1646 AGACTGCAAGCGCAGAGATCAGGGAATTTGTACAACAGGTGAAGCAGAGCAGAGAAAT 1705
 Db 1752 AGACTGCAAGCGCAGAGATCAGGGAATTTGTACAACAGGTGAAGCAGAGCAGAGAAAT 1811
 QY 1706 ACCCTGAGGAGTGCCTCAACAGCGCAAGAGTGAAGCTCCACAGGACCAAGCAAAACA 1765

Db 1812 ACCCTGAGGAGTGCCTCAACAGCGCAGAGCAAGTGAAGCTCCCAAGGCAACCAAAACA 1871
 QY 1766 ACGGATGATATGAGCCCTTCCAAACCTGACAGAAATGAGACCAACGCAAGCCAGCAGAT 1825
 Db 1872 ACGGATGATATGAGCCCTTCCAAACCTGACAGAAATGAGACCAACGCAAGCCAGCAGAT 1931
 QY 1826 CGGAGCAAAACCAAGCAATCTGAGGAATGAGAAATCTGGGAGGCGGCGCAAGGACTCT 1885
 Db 1932 CGGAGCAAAACCAAGCAATCTGAGGAATGAGAAATCTGGGAGGCGGCGCAAGGACTCT 1991
 QY 1886 GCCGAGGCCCTGAGAACCC 1904
 Db 1992 GCCGAGGCCCTGAGAACCC 2010

Search completed: July 24, 2004, 13:04:44
 Job time : 12667.7 secs

This Page Blank (uspto)


```
Db      2041 CGGTATCCCTTTAGTGAACATPACATGAGTGAGCGTGTCAAGCCAAAGCAAAATGCAC 2100
QY      2101 ACCCTTTTCTGTGGCAATCGTCTGTACATGTGTATCATATTTAGAAAGGAAAGATG 2160
Db      2101 ACCCTTTTCTGTGGCAATCGTCTGTACATGTGTATCATATTTAGAAAGGAAAGATG 2160
QY      2161 TTAAAGATATGTGGCTGTGGTTACAACAGGTGCTGACAGCGTAATATATTTAGAAAT 2220
Db      2161 TTAAAGATATGTGGCTGTGGTTACAACAGGTGCTGACAGCGTAATATTTAGAAAT 2220
QY      2221 AATATATCAATATCTCAACTCACTCAATTTTATCATATTTATTTATTTTCTTT 2280
Db      2221 AATATATCAATATCTCAACTCACTCAATTTTATCATATTTATTTATTTTCTTT 2280
QY      2281 TTAAGAGAAAGCAGGCTTTTGTAGACTTTAAAGAAATTAAGCTTTGGAGGTCTCAGG 2340
Db      2281 TTAAGAGAAAGCAGGCTTTTGTAGACTTTAAAGAAATTAAGCTTTGGAGGTCTCAGG 2340
QY      2341 TGTAGAGAGAGCTTTGAGGCCACCCGACAAATTTCAACCAGAGGAAATCTCGTGGGA 2400
Db      2341 TGTAGAGAGAGCTTTGAGGCCACCCGACAAATTTCAACCAGAGGAAATCTCGTGGGA 2400
QY      2401 AGGACACTCAGCGGAGCTTGTGATACCTGTGTATGTCAACAGAGGATACCGTCTCT 2460
Db      2401 AGGACACTCAGCGGAGCTTGTGATACCTGTGTATGTCAACAGAGGATACCGTCTCT 2460
QY      2461 TGAAGAGGAACTCTGTCACTCCTCATGCTGTAGCTCATACCCATTTCTTTGCT 2520
Db      2461 TGAAGAGGAACTCTGTCACTCCTCATGCTGTAGCTCATACCCATTTCTTTGCT 2520
QY      2521 TTCAAGGTTTAAACTGGTTTTTGTGATCTGTATTAATTTCTGTCTCTCTGT 2580
Db      2521 TTCAAGGTTTAAACTGGTTTTTGTGATCTGTATTAATTTCTGTCTCTCTGT 2580
QY      2581 TATCTTCCCTCCCTCCCTCCCTCCCTCTTCTCATCTCATCTCATCTTTGAAATTCCTCAT 2640
Db      2581 TATCTTCCCTCCCTCCCTCCCTCCCTCTTCTCATCTCATCTCATCTTTGAAATTCCTCAT 2640
QY      2641 CCCCTCATCTCAATCCCGTATCTAGGACACCCCGCCCGCCAGGCAAGCGTGTCTGA 2700
Db      2641 CCCCTCATCTCAATCCCGTATCTAGGACACCCCGCCCGCCAGGCAAGCGTGTCTGA 2700
QY      2701 GTATCATATCAACAAAGGAAAGCAAGCAACAAACAGCTCAACTTACACTT 2760
Db      2701 GTATCATATCAACAAAGGAAAGCAAGCAACAAACAGCTCAACTTACACTT 2760
QY      2761 GGTACTCTAAAGAAAGAGTCAATGTACTTGTCTTCTGAGCTTTGAAAGAGAAACA 2820
Db      2761 GGTACTCTAAAGAAAGAGTCAATGTACTTGTCTTCTGAGCTTTGAAAGAGAAACA 2820
QY      2821 GGAACCCACCAACCAACCAATCAACCAACCAAGAAATTTCCAAATGAAAGATGT 2880
Db      2821 GGAACCCACCAACCAACCAATCAACCAACCAAGAAATTTCCAAATGAAAGATGT 2880
QY      2881 AATTTGTCTTTTGCATTTTGTGTATAGCCATCATATTTACGAAATGATTTCTTTC 2940
Db      2881 AATTTGTCTTTTGCATTTTGTGTATAGCCATCATATTTACGAAATGATTTCTTTC 2940
QY      2941 TTTAAAGAAAGAAATGTGGAGAAAGTAAATTTTCAAGGTGTTGGCCAGGGCGTT 3000
Db      2941 TTTAAAGAAAGAAATGTGGAGAAAGTAAATTTTCAAGGTGTTGGCCAGGGCGTT 3000
QY      3001 AAATTCACAGATTTTAAACGAGAAACACACAGAGAGTACTCTCGGTGTTT 3060
Db      3001 AAATTCACAGATTTTAAACGAGAAACACACAGAGAGTACTCTCGGTGTTT 3060
QY      3061 CCTCAGACCTTGTCTGTGTCTTGTAGATTTTGAAGTGTATGTTGAGAGAT 3120
Db      3061 CCTCAGACCTTGTCTGTGTCTTGTAGATTTTGAAGTGTATGTTGAGAGAT 3120
QY      3121 TTTTATTTTATTTTAAATGAGTTGAGAAATTAAGATCAACTGCGACGCTG 3180
Db      3121 TTTTATTTTATTTTAAATGAGTTGAGAAATTAAGATCAACTGCGACGCTG 3180
```

```
QY      3181 GAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTCTGAAATTTCTTCGCTAGCCAGAA 3240
Db      3181 GAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTCTGAAATTTCTTCGCTAGCCAGAA 3240
QY      3241 CCNATATGCGCTTTTGGACAAACCTTGAAATGTTATTT 3283
Db      3241 CCNATATGCGCTTTTGGACAAACCTTGAAATGTTATTT 3283

RESULT 2
ID      AAS70982 standard; cDNA, 3667 BP.
XX
XX      AAS70982;
AC
XX
XX      13-FEB-2002 (first entry)
DT
XX
XX      DNA encoding novel human diagnostic protein #6786.
DE
XX
XX      Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW      food supplement; medical imaging; diagnostic; genetic disorder; ss.
OS
XX      Homo sapiens.
EN      WO200175067-A2.
XX
XX      11-OCT-2001.
PD
XX
XX      30-MAR-2001; 2001WO-US008631.
PF
XX
XX      31-MAR-2000; 2000US-00540217.
PR      23-AUG-2000; 2000US-00649167.
XX
XX      (HXS-) HXSEQ INC.
PA
XX      Dmanac RT, Liu C, Tang YT,
PI
XX      WPI; 2001-639362/73.
DR      P-PSDB; ABG06795.
XX
XX      New isolated polynucleotide and encoded polypeptides, useful in
PT      diagnostics, forensics, gene mapping, identification of mutations
PT      responsible for genetic disorders or other traits and to assess
PT      biodiversity.
XX
XX      Claim 1; SEQ ID NO 6786; 103pp; English.
XX
XX      The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC      sequences. (I) is useful as hybridisation probes, polymerase chain
CC      reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC      and in recombinant production of (II). The polynucleotides are also used
CC      in diagnostics as expressed sequence tags for identifying expressed
CC      genes. (I) is useful in gene therapy techniques to restore normal
CC      activity of (II) or to treat disease states involving (II). (II) is
CC      useful for generating antibodies against it, detecting or quantitating a
CC      polypeptide in tissue, as molecular weight markers and as a food
CC      supplement. (II) and its binding partners are useful in medical imaging
CC      of sites expressing (II). (I) and (II) are useful for treating disorders
CC      involving aberrant protein expression or biological activity. The
CC      polypeptide and polynucleotide sequences have applications in
CC      diagnostics, forensics, gene mapping, identification of mutations
CC      responsible for genetic disorders or other traits to assess biodiversity
CC      and to produce other types of data and products dependent on DNA and
CC      amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC      coding sequences of the invention. Note: The sequence data for this
CC      patent did not appear in the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences
XX
XX      Sequence 3667 BP; 1081 A; 882 C; 846 G; 858 T; 0 U; 0 Other;
SQ
Query Match      99.1%; Score 3253.8; DB 5; Length 3667;
```

	Best Local Similarity	99.8%	Pred: No. 0;	Mismatches 3277;	Conservative 0;	Mismatches 3;	Indels 2;	Gaps 2;
QY	3	CAGCGGAGGAGGGCAGAGAGCGCCGGGTATCCGGGCCGGGGAGCCCGGGCTTCCGGGAA	62					
Db	369	CAACGAGGAGCGAGGAGCGCGGGGTACCGGGCCGGGGAGACCGGGGCTTCGGGGAA	428					
QY	63	GAGACGGATATGAACAAGCTTTAATCATCGGGAAACTGAGCCCCGCCTCACCGCCGAGA	122					
Db	429	GAGCGATATATGAACAAGCTTTAATCATCGGGAACTGAGCCCCGCCTCACCGCCGAGA	488					
QY	123	CCTCCGGCAGCTCTTTGGGGACAAGAACCTCCCTGGCGGACAGTCTTGTAATGC	182					
Db	489	CTCCGGCAGCTCTTTGGGGACAAGAACCTGCCCTGGCGGACAGTCTTGTAATGC	548					
QY	183	CGGCTAAGCCTTGCTGGACTATCCCCACAGAACTGGGGCATTCGGGCATGAGACCCCT	242					
Db	549	CGGCTAAGCCTTGCTGGACTATCCCCACAGAACTGGGGCATTCGGGCATGAGACCCCT	608					
QY	243	CTCGGGTAAATGGAAATGCTAGGGAAATCATAGAAATTGATTACTAGTCTTAAAAA	302					
Db	609	CTCGGGTAAATGGAAATGCTAGGGAAATCATAGAAATTGATTACTAGTCTTAAAAA	668					
QY	303	GCTAAGAGCAGGAAATTCAGATTGAAAATCTCCTCTCACCTGCAGTGGGAGGCTT	362					
Db	669	GCTAAGAGCAGGAAATTCAGATTGAAAATCTCCTCTCACCTGCAGTGGGAGGCTT	728					
QY	363	GGATGCACTTTTGGCTCAATATGGGACAGTGGAGATGTGAAACAAGTCAACACAGAC	422					
Db	729	GGATGCACTTTTGGCTCAATATGGGACAGTGGAGATGTGAAACAAGTCAACACAGAC	788					
QY	423	AGAAACCCTCGTTGTCAAAGTCAATATGCAACAAAGAAAGAAATAATAGCTATGA	482					
Db	789	AGAAACCCTCGTTGTCAAAGTCAATATGCAACAAAGAAAGAAATAATAGCTATGA	848					
QY	483	GAGCTAAGCGGGGACATCATGTTGAGAACTACTCCTTCAAGATTCTCAATGCCGAGTA	542					
Db	849	GAGCTAAGCGGGGACATCATGTTGAGAACTACTCCTTCAAGATTCTCAATGCCGAGTA	908					
QY	543	AAGGTGAGCTCCCTTGCGCCCTCAAGCGAGCCAGGTGGGGACCACTTTCGGGGA	602					
Db	909	AAGGTGAGCTCCCTTGCGCCCTCAAGCGAGCCAGGTGGGGACCACTTTCGGGGA	968					
QY	603	GCAAGGCACGCCCCCTGGGGGACATTTCTCAAGCCACAGATTTATTTCCCCTGGGAGT	662					
Db	969	GCAAGGCACGCCCCCTGGGGGACATTTCTCAAGCCACAGATTTATTTCCCCTGGGAGT	1028					
QY	663	CCCTGGCCCCACCGATTTGTGTGGCATCATCGGAAAGAGGGCTTGACATTAAGAA	722					
Db	1029	CCCTGGCCCCACCGATTTGTGTGGCATCATCGGAAAGAGGGCTTGACATTAAGAA	1088					
QY	723	CATCACTMAAGCAGACCAGTCCCGGGTAGATATCCATAGAAAAGAGAACTGTGAGCTGC	782					
Db	1089	CATCACTMAAGCAGACCAGTCCCGGGTAGATATCCATAGAAAAGAGAACTGTGAGCTGC	1148					
QY	783	AGAGAAAGCTGTACCATTCATATGCAACCCCAAGGGGACTTGTAGAGCATGCCATGAT	842					
Db	1149	AGAGAAAGCTGTACCATTCATATGCAACCCCAAGGGGACTTGTAGAGCATGCCATGAT	1208					
QY	843	TCTTGAATCATGCAAGAAAGAGGCAAGATAGACAAACTAGCCGAAGATTTCTCTGAA	902					
Db	1209	TCTTGAATCATGCAAGAAAGAGGCAAGATAGACAAACTAGCCGAAGATTTCTCTGAA	1268					
QY	903	AATCTTGACACACATAGCTTGTGTGGAAGACTGATGTGAAAAGAAAGGACAAATTGAA	962					
Db	1269	AATCTTGACACACATAGCTTGTGTGGAAGACTGATGTGAAAAGAAAGGACAAATTGAA	1328					
QY	963	GAAATTTGAACATGAAGACAGGGACCAAGATTAACAATCTCATCTTTGACAGATTTGAGCAT	1022					
Db	1329	GAAATTTGAACATGAAGACAGGGACCAAGATTAACAATCTCATCTTTGACAGATTTGAGCAT	1388					
QY	1023	ATACACCCCGAAAGAACATCACTGTGAAGGACAGTTGAGGCTGTGCCATGTCTGA	1082					

Db	1389	ATCACACC	GGAAAGAA	ACATCACT	GTGAAGG	CA	CAGTTG	AGGCGCTG	TCCAGT	CTGA	1448				
Qy	1083	GATTAAGA	TATTAA	GAAGCT	GTGTGA	GGCCTTTG	AAAAAT	GATATG	TGTGGC	CTTTAA	CAC	1142			
Db	1449	GATTAGAT	TATTAG	AAGAACT	GTGTGA	GGCCTTTG	AAAAAT	GATATG	TGTGGC	CTTTAA	CAC	1508			
Qy	1143	CCATCTCC	GGATAT	CTTCCAG	CTGTGA	AGCCCTCA	CCAGTTTG	GGCCCTGT	CCCGAC	TA	1202				
Db	1509	CCACTCCG	ATATCTT	CCAGCT	TGAACCC	CA	CCAGTTTG	GGCCCTGT	CCCGAC	TA	1568				
Qy	1203	TCACTCTT	ATCCAG	ACAGAGAT	TGTGAAT	CTTTAT	CCCAAC	CCAGCTGT	GGCGC		1262				
Db	1569	TCACTCTT	ATCCAG	ACAGAGAT	TGTGAAT	CTTTAT	CCCAAC	CCAGCTGT	GGCGC		1628				
Qy	1263	CATCAT	TGGGAA	GAAGGGG	CA	CATCAAA	CAAGCTTG	GGCGAGAT	TGCGCG	AGCTCTAT	1322				
Db	1629	CATCAT	TGGGAA	GAAGGGG	CA	CATCAAA	CAAGCTTG	GGCGAGAT	TGCGCG	AGCTCTAT	1688				
Qy	1323	CAAGAT	TGCCCC	TGGGAA	GGCCCA	GCCTCAG	CGAAAGAT	TGTTCAT	CATCAC	CGGGCC	1382				
Db	1689	CAAGAT	TGCCCC	TGGGAA	GGCCCA	GCCTCAG	CGAAAGAT	TGTTCAT	CATCAC	CGGGCC	1748				
Qy	1383	ACCGAAGC	CCAGTT	CAAGGCC	CA	GGGACGGAT	CTTTGG	AACTGA	AAAGAA	CTT	1442				
Db	1749	ACCGAAGC	CCAGTT	CAAGGCC	CA	GGGACGGAT	CTTTGG	AACTGA	AAAGAA	CTT	1808				
Qy	1443	CTTTAA	CCCCAA	GAAGAGT	GAAGCT	GAAGAG	CGCATAT	CAGATG	CGCCTT	CCAC	AGC	1502			
Db	1809	CTTTAA	CCCCAA	GAAGAGT	GAAGCT	GAAGAG	CGCATAT	CAGATG	CGCCTT	CCAC	AGC	1868			
Qy	1503	TGGCCGGG	GTATTT	GGCAAA	GGTGG	CAAGACCG	GTAG	AGAACTG	CGAGAA	CTTAA	CGTGC	1562			
Db	1869	TGGCCGGG	GTATTT	GGCAAA	GGTGG	CAAGACCG	GTAG	AGAACTG	CGAGAA	CTTAA	CGTGC	1928			
Qy	1563	AGAAGT	CATCGT	GCCTGT	GA	CCAA	CGCCAG	ATGA	AAATGA	GGAA	GTGTCAGAT	1622			
Db	1929	AGAAGT	CATCGT	GCCTGT	GA	CCAA	CGCCAG	ATGA	AAATGA	GGAA	GTGTCAGAT	1988			
Qy	1623	TATGGG	CACTTCTT	TGCTAG	CCGAC	CTGCA	CAGGCA	GAATCG	AGGAAAT	TGTACA	CA	1682			
Db	1989	TATGGG	CACTTCTT	TGCTAG	CCGAC	CTGCA	CAGGCA	GAATCG	AGGAAAT	TGTACA	CA	2048			
Qy	1683	GGTGAAC	CAGAGG	AGCA	GAATAC	CCCTCA	GGAGT	CGCTC	CA	CGCAG	CGACGA	GTGAG	1742		
Db	2049	GGTGAAC	CAGAGG	AGCA	GAATAC	CCCTCA	GGAGT	CGCTC	CA	CGCAG	CGACGA	GTGAG	2108		
Qy	1743	CTCCCA	CAGGCA	CCAGCA	AAACAA	CGATGA	TGTAG	CCCTT	CCAA	CACTTGA	CAGAA	TG	1802		
Db	2109	CTCCCA	CAGGCA	CCAGCA	AAACAA	CGATGA	TGTAG	CCCTT	CCAA	CACTTGA	CAGAA	TG	2168		
Qy	1803	AGAC	CCAA	CGCAG	CCAG	CCAGT	TGGGAG	CAAA	CCAA	AGACAT	CTGA	GAATGA	GAAGT	1862	
Db	2169	AGAC	CCAA	CGCAG	CCAG	CCAGT	TGGGAG	CAAA	CCAA	AGACAT	CTGA	GAATGA	GAAGT	2228	
Qy	1863	CTGGGAG	AGCG	CCAG	CACTTGC	CGAG	GCCTTGA	AA	CCCCAG	GGGCG	CGAGAG	GGGCG	1922		
Db	2229	CTGGGAG	AGCG	CCAG	CACTTGC	CGAG	GCCTTGA	AA	CCCCAG	GGGCG	CGAGAG	GGGCG	2288		
Qy	1923	GGGGAAG	GTAC	CCAG	GTTTG	CCAG	ACCA	CCGAG	CCCCCG	CTCCCCCG	CCCA	GGGCTT	1982		
Db	2289	GGGGAAG	GTAC	CCAG	GTTTG	CCAG	ACCA	CCGAG	CCCCCG	CTCCCCCG	CCCA	GGGCTT	2348		
Qy	1983	CTGCA	GGCTTCA	GGC	CATCTCA	CTTCA	CCACT	CGATCT	CTCTG	AACTCC	CA	CGACG	2042		
Db	2349	CTGCA	GGCTTCA	GGC	CATCTCA	CTTCA	CCACT	CGATCT	CTCTG	AACTCC	CA	CGACG	2408		
Qy	2043	CTATCC	CTTTAG	TGA	ACTA	CATAG	TGA	AGTGA	AGTGT	TGA	AGGAA	TG	CA	2102	
Db	2409	CTATCC	CTTTAG	TGA	ACTA	CATAG	TGA	AGTGA	AGTGT	TGA	AGGAA	TG	CA	2468	
Qy	2103	CTTTTTC	TGTG	GA	CAATG	CTCTCT	GTAC	TGTGT	GTAC	TATTTA	GA	AGGGA	GAT	TT	2162
Db	2469	CTTTTTC	TGTG	GA	CAATG	CTCTCT	GTAC	TGTGT	GTAC	TATTTA	GA	AGGGA	GAT	TT	2528

QY 2163 AAGATATGTGGCTGTGGGTTACACAGGGTGCTGACGGGTAATATTTTGAATAA 2222
DB 2529 AAGATATGTGGCTGTGGGTTACACAGGGTGCTGACGGGTAATATTTTGAATAA 2588
QY 2223 TATATCAAAATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2282
DB 2589 TATATCAAAATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 2648
QY 2283 AAGAGAAAGCAAGGCTTTTCTAGACTTTTAAAGATAAAGCTTTTGGAGGCTCACGGTG 2342
DB 2649 AAGAGAAAGCAAGGCTTTTCTAGACTTTTAAAGATAAAGCTTTTGGAGGCTCACGGTG 2708
QY 2343 TAGAGAGAGCTTTAGAGGCAACCCGACAAATTCACCCAGAGGGAATCTGTGGAG 2402
DB 2709 TAGAGAGAGCTTTAGAGGCAACCCGACAAATTCACCCAGAGGGAATCTGTGGAG 2768
QY 2403 GACACTCAGCGAGTTCGATGATCAGCTGTGTATGTCAACAGAGGGAATCCGTCTCTG 2462
DB 2769 GACACTCAGCGAGTTCGATGATCAGCTGTGTATGTCAACAGAGGGAATCCGTCTCTG 2828
QY 2463 AAGAGGAACCTGTGACCTCCATAGCTGTGTACTCAATACCACTTTCTCTTGTG 2522
DB 2829 AAGAGGAACCTGTGACCTCCATAGCTGTGTACTCAATACCACTTTCTCTTGTG 2888
QY 2523 CACAGGTTTAACTGTTTTTGTGATGATGATGATGATGATGATGATGATGATGAT 2582
DB 2889 CACAGGTTTAACTGTTTTTGTGATGATGATGATGATGATGATGATGATGATGAT 2948
QY 2583 TCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 2642
DB 2949 TCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3008
QY 2643 CTCATCTCCATCCCGTATCTACGACCCCGCCCGCCAGGCAAGAGTCTGTAGT 2702
DB 3009 CTCATCTCCATCCCGTATCTACGACCCCGCCCGCCAGGCAAGAGTCTGTAGT 3067
QY 2703 ATCAATCACAACAAAGGAGCAAAAGGAGCAACACAAAGGCTCACTTACCTTGG 2762
DB 3068 ATCAATCACAACAAAGGAGCAAAAGGAGCAACACAAAGGCTCACTTACCTTGG 3127
QY 2763 TTACTCAAAAGACAGAGTCAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 2822
DB 3128 TTACTCAAAAGACAGAGTCAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 3187
QY 2823 AACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAATTCACCAATGAAATGTA 2882
DB 3188 AACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAATTCACCAATGAAATGTA 3247
QY 2883 TTTGTCTTTTGTGATTTGTGTATTAAGGCTCAATTTTCAAGGTAATGTAATGTAATGTA 2942
DB 3248 TTTGTCTTTTGTGATTTGTGTATTAAGGCTCAATTTTCAAGGTAATGTAATGTAATGTA 3307
QY 2943 T-AAAAAAAAAATGTGAGGAAAGTAAGAAATTTTACCAAGTTGTTGAGGCGGCTTA 3001
DB 3308 TAAAAAATAAATGTGAGGAAAGTAAGAAATTTTACCAAGTTGTTGAGGCGGCTTA 3367
QY 3002 AATTCACAGTTTTAAAGAGAAACACACAGAGAAAGTCACTGAGGTGTTTATAC 3061
DB 3368 AATTCACAGTTTTAAAGAGAAACACACAGAGAAAGTCACTGAGGTGTTTATAC 3427
QY 3062 CTCAGACCTTGTGCTGTGTTCCCTTAGAGATTTGTAAAGTGTAGTGTAGAGATT 3121
DB 3428 CTCAGACCTTGTGCTGTGTTCCCTTAGAGATTTGTAAAGTGTAGTGTAGAGATT 3487
QY 3122 TTTTATTTTTTAAATAAATGATGAAATTAAGATATCACTGAGCGCTGG 3181
DB 3488 TTTTATTTTTTAAATAAATGATGAAATTAAGATATCACTGAGCGCTGG 3547
QY 3182 AGAAGGTGACGTCCAGTGTGACACGCTGTTCTGAATGTCTTCCGTAGCCAGAAC 3241
DB 3548 AGAAGGTGACGTCCAGTGTGACACGCTGTTCTGAATGTCTTCCGTAGCCAGAAC 3607

QY 3242 CNAATAGGCTTTCTTTGGACAAACCTGAAATGTTATTT 3283
DB 3608 CTATATGGCTTTCTTTGGACAAACCTGAAATGTTATTT 3649
RESULT 3
AB876442
ID AB876442 standard, cDNA, 3667 BP.
XX
XX AB876442;
XX
DT 11-DEC-2002 (first entry)
XX
DE cDNA encoding human ovarian cancer marker M452.
XX
XX Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nonderculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200271928-A2.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US007826.
XX
PR 14-MAR-2001; 2001US-0276025P.
PR 14-MAR-2001; 2001US-0276026P.
PR 10-AUG-2001; 2001US-0311732P.
PR 19-SEP-2001; 2001US-0323580P.
PR 26-SEP-2001; 2001US-0324967P.
PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG,
PI Meyers RE, Morrisey ME, Olandt PU, Sen A, Vleby PO, Mills GB,
PI Baet RC, Lu K, Schmandt RE, Zhao X, Ghatt K;
XX
XX WPI: 2002-723277/78.
XX
XX P-PSDB; ABG96346.
XX
XX
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.
XX
XX
XX Disclosure; Page 262-263; 481pp; English.
XX
XX The present invention relates to a new method for assessing whether a
XX patient is afflicted with ovarian cancer. The method involves comparing
XX the expression level of a marker in a patient sample and the normal level
XX of expression of the marker in a control non-ovarian cancer sample, where
XX the marker is selected from 363 cancer markers described in the
XX specification. The method of the invention is useful in diagnosing or
XX characterizing cancer, in detecting the presence of cancer as early as
XX possible, and the recurrence of ovarian cancer. The method may also be of
XX particular use with patients having an enhanced risk of developing
XX ovarian cancer (e.g. patients having a familial history of ovarian
XX cancer). The cancer markers may be used in the management and treatment
XX of e.g. brain and central nervous system disorders (e.g. bacterial and
XX viral meningitis, Alzheimer's disease or Parkinson's disease), brain
XX disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
XX inflammations (e.g. bacterial or viral meningitis or encephalitis),
XX testicular disorders (e.g. nonderculous granulomatous orchitis),
XX connective tissue disorders, or heart disorders (e.g. ischaemic heart
XX disease or atherosclerosis). The compositions and methods may also be

used in assessing the histological type of neoplasm associated with ovarian cancer, monitoring the progression of ovarian cancer, determining whether ovarian cancer has metastasized or is likely to metastasize, selecting a composition for inhibiting ovarian cancer, assessing the ovarian carcinogenic potential of a compound, or inhibiting ovarian cancer or at risk of developing ovarian cancer. The present nucleic acid sequence encodes one of the ovarian cancer markers described in the invention

Sequence 3667 BP; 1081 A; 882 C; 846 G; 858 T; 0 U; 0 Other;

Query Match 99.1%; Score 3253.8; DB 6; Length 3667;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3277; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

```
3 CAGCGGAGAGGCGGAGGAGCGCGGGTACCGGGCCGGGGAGACCGGGCTCTCGGGGA 62
Db CAAAGGAGAGGCGGAGGAGCGCGGGTACCGGGCCGGGGAGACCGGGCTCTCGGGGA 428
Qy GAGAGGATGATGAACAACCTTTACATGGGGAACCTGAGCCCGCGCTCAGCGCGAGCA 122
Db GAGAGGATGATGAACAACCTTTACATGGGGAACCTGAGCCCGCGCTCAGCGCGAGCA 488
Qy GAGAGGATGATGAACAACCTTTACATGGGGAACCTGAGCCCGCGCTCAGCGCGAGCA 182
Db CTTCCGGCAGCTCTTTGGGGAAGAGAGCTGCCCTGGCGGAGCAGGTCTGCTGAAGTC 548
Qy CGGCTACGCTCTTGTGAGCTACCCCGAGCACTGGGCAATCGGGCAATGAGAGCCCT 242
Db CGGCTACGCTCTTGTGAGCTACCCCGAGCACTGGGCAATCGGGCAATGAGAGCCCT 608
Qy CGGCTACGCTCTTGTGAGCTACCCCGAGCACTGGGCAATCGGGCAATGAGAGCCCT 243
Db CTTCCGGTAAAGTGAATTCATGGGGAATTCATGAAATGATTAATCTCACTCTTAAAA 302
Qy CTTCCGGTAAAGTGAATTCATGGGGAATTCATGAAATGATTAATCTCACTCTTAAAA 668
Db CTTCCGGTAAAGTGAATTCATGGGGAATTCATGAAATGATTAATCTCACTCTTAAAA 303
Qy GCTAAAGAGCAGGAAATTCAGATTCGAAACATCCCTCTCACTCAGTGGAGGTGT 362
Db GCTAAAGAGCAGGAAATTCAGATTCGAAACATCCCTCTCACTCAGTGGAGGTGT 728
Qy GCTAAAGAGCAGGAAATTCAGATTCGAAACATCCCTCTCACTCAGTGGAGGTGT 363
Db GCTAAAGAGCAGGAAATTCAGATTCGAAACATCCCTCTCACTCAGTGGAGGTGT 728
Qy GATGAGATCTTTGGCTCAATATGAGGAGATGAGGATGAGGACCAATCAACAGAC 422
Db GATGAGATCTTTGGCTCAATATGAGGAGATGAGGATGAGGACCAATCAACAGAC 788
Qy GATGAGATCTTTGGCTCAATATGAGGAGATGAGGATGAGGACCAATCAACAGAC 423
Db GATGAGATCTTTGGCTCAATATGAGGAGATGAGGATGAGGACCAATCAACAGAC 789
Qy AGAAACCGCGCTGTCAATATGAGGAGATGAGGAGATGAGGACCAATCAACAGAC 482
Db AGAAACCGCGCTGTCAATATGAGGAGATGAGGAGATGAGGACCAATCAACAGAC 848
Qy AGAAACCGCGCTGTCAATATGAGGAGATGAGGAGATGAGGACCAATCAACAGAC 483
Db AGAAACCGCGCTGTCAATATGAGGAGATGAGGAGATGAGGACCAATCAACAGAC 849
Qy GAAAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 542
Db GAAAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 908
Qy GAAAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 543
Db GAAAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 909
Qy AGAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 968
Db AGAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 969
Qy GAAAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 603
Db GAAAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 1028
Qy GAAAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 663
Db GAAAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 1029
Qy CTTGAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 723
Db CTTGAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 1089
Qy CTTGAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 783
Db CTTGAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 1149
Qy AGGAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 843
Db AGGAGTAAAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGATGAGGAGAT 902
```

```
1209 TCTTGAATCATGAGAAAGGAGATGAGACCAATCTAGCCGAGAGATTCCTCTGA 1268
Qy AATCTGGCACAATGAGCTGTGTAAGATGATGAAAGAGAGAGAGAGATTTGA 962
Db AATCTGGCACAATGAGCTGTGTAAGATGATGAAAGAGAGAGAGAGATTTGA 1328
Qy AATCTGGCACAATGAGCTGTGTAAGATGATGAAAGAGAGAGAGAGATTTGA 1022
Db AATCTGGCACAATGAGCTGTGTAAGATGATGAAAGAGAGAGAGAGATTTGA 1388
Qy GAAATTTGAATGAGAGAGAGAGAGAGATGAGAGATGAGAGATGAGAGATTTGA 1082
Db GAAATTTGAATGAGAGAGAGAGAGAGATGAGAGATGAGAGATGAGAGATTTGA 1448
Qy AATCAACCCGAGAGAGAGAGATGAGAGAGAGAGAGATGAGAGAGATTTGA 1449
Db AATCAACCCGAGAGAGAGAGATGAGAGAGAGAGAGATGAGAGAGATTTGA 1508
Qy GATAGAGATTTGAAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1262
Db GATAGAGATTTGAAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1322
Qy TCACTCTTATCCAGAGAGAGAGATGAGATCTTTCAATCCCAACAGAGCTGTGAGC 1628
Db TCACTCTTATCCAGAGAGAGAGATGAGATCTTTCAATCCCAACAGAGCTGTGAGC 1688
Qy CATCATCGGAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1382
Db CATCATCGGAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1442
Qy CAAAGTTCGCTGTGAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1748
Db CAAAGTTCGCTGTGAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1808
Qy ACGGAGAGAGAGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1502
Db ACGGAGAGAGAGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1562
Qy CTTTAAACCCCAAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1622
Db CTTTAAACCCCAAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGATTTGA 1688
Qy TGGCCGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAGAGAGATTTGA 1868
Db TGGCCGGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAGAGAGATTTGA 1928
Qy AGAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAGAGAGATTTGA 1622
Db AGAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAGAGAGATTTGA 1688
Qy TATCGGAGATTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 1682
Db TATCGGAGATTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 1742
Qy TATCGGAGATTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 2048
Db TATCGGAGATTTGCTGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 2108
Qy GGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 1802
Db GGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 1862
Qy GGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 2168
Db GGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 2228
Qy GGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 2288
Db GGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTGA 2348
```


Db 1 GGCACGAGAGAGGAGGAGCGCCGGTACCGGGCCGGGGAGCCGGGGCTTCGGGG 60
QY 61 AAGAGACGATGATGAAACAAGCTTTACATCGGGAACCTGAGCCCGGCGCTCACCCGAC 120
Db 61 AAGAGACGATGATGAAACAAGCTTTACATCGGGAACCTGAGCCCGGCGCTCACCCGAC 120
QY 121 GACCTCCGGCAGCTCTTTGGGGGACAGGAAGCTGCCCTGGCGGACAGGCTCTGCTGAAG 180
Db 121 GACCTCCGGCAGCTCTTTGGGGGACAGGAAGCTGCCCTGGCGGACAGGCTCTGCTGAAG 180
QY 181 TCCGGCTACGCTTCCTGGAGCTACCCGACCAAGATCGGGCCATCCGAGAC 240
Db 181 TCCGGCTACGCTTCCTGGAGCTACCCGACCAAGATCGGGCCATCCGAGAC 240
QY 241 CTCTCGGGTAAAGTGAATTTGCAATGGGAAATCATGAAAGTTGATTAATCACTCTTAA 300
Db 241 CTCTCGGGTAAAGTGAATTTGCAATGGGAAATCATGAAAGTTGATTAATCACTCTTAA 300
QY 301 AAGCTAAGAGAGAGGAAATTCAGATTGMAACATCCCTCCTCAGCTGAGTGGAGTG 360
Db 301 AAGCTAAGAGAGGAAATTCAGATTGMAACATCCCTCCTCAGCTGAGTGGAGTG 360
QY 361 TTGGATGACTTTTGGCTCAATATGGGACAGTGAATGTGAAACAGTCAACACAGAC 420
Db 361 TTGGATGACTTTTGGCTCAATATGGGACAGTGAATGTGAAACAGTCAACACAGAC 420
QY 421 ACAGAAACCGCGCTTGTCAAGTCAATATGCAACAGAGAAAGACCAAAATAGCCATG 480
Db 421 ACAGAAACCGCGCTTGTCAAGTCAATATGCAACAGAGAAAGACCAAAATAGCCATG 480
QY 481 GAGAAGCTAAGCGGAGTCAAGTTTGAAGCTACTCCTTCAAGATTTCTTACATCCCGAT 540
Db 481 GAGAAGCTAAGCGGAGTCAAGTTTGAAGCTACTCCTTCAAGATTTCTTACATCCCGAT 540
QY 541 GAAGAGTGAAGTCCCTTTCGCCCTCAGCGAGCCGAGGTGGGACCACTCTTCCCG 600
Db 541 GAAGAGTGAAGTCCCTTTCGCCCTCAGCGAGCCGAGGTGGGACCACTCTTCCCG 600
QY 601 GAGCAAGGCGAGCCCTTGGGGGACCTTCTCAAGCCAGACAGATTAATTCGCGCTGG 660
Db 601 GAGCAAGGCGAGCCCTTGGGGGACCTTCTCAAGCCAGACAGATTAATTCGCGCTGG 660
QY 661 ATCTGCTCCCAACCCAGTTTGTGGTGCATCATCGAAAGAGGCTTGAACATTAAG 720
Db 661 ATCTGCTCCCAACCCAGTTTGTGGTGCATCATCGAAAGAGGCTTGAACATTAAG 720
QY 721 AACATCACTAAGACAGCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGAGCT 780
Db 721 AACATCACTAAGACAGCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGAGCT 780
QY 781 GCGAGAGGCTGTCAACATTCATGCAACCCAGAGGGGACTTCTGAAGATGCCGATG 840
Db 781 GCGAGAGGCTGTCAACATTCATGCAACCCAGAGGGGACTTCTGAAGATGCCGATG 840
QY 841 ATTCTTGAATCATGCGAAAGAGGAGAGTGAAGCAAACTAGCGAAGAGATTCCTCTG 900
Db 841 ATTCTTGAATCATGCGAAAGAGGAGAGTGAAGCAAACTAGCGAAGAGATTCCTCTG 900
QY 901 AAAATCTTGGCAGCAATGCTGTGGTGGAAAGCTGATGGAAAGAGGAGAGAAATTTG 960
Db 901 AAAATCTTGGCAGCAATGCTGTGGTGGAAAGCTGATGGAAAGAGGAGAGAAATTTG 960
QY 961 AAGAAATTTGAACATGAAACAGGAGCAAGATTAACATTCATTTTGGCAGGATTTAGC 1020
Db 961 AAGAAATTTGAACATGAAACAGGAGCAAGATTAACATTCATTTTGGCAGGATTTAGC 1020
QY 1021 ATATACACCCGAGAAAGACCATCATGTGAAGGAGCAAGTTGAGGCTGTGCCAGTCT 1080
Db 1021 ATATACACCCGAGAAAGACCATCATGTGAAGGAGCAAGTTGAGGCTGTGCCAGTCT 1080
QY 1081 GAGATAGAGTTATGAAGAGCTGCTGAGGCTTTGAAATGATATGCTGGCTTAAAC 1140
Db 1081 GAGATAGAGTTATGAAGAGCTGCTGAGGCTTTGAAATGATATGCTGGCTTAAAC 1140

QY 1141 ----- 1140
Db 1141 CAACAAGCCATCTGATCCAGGGTTGAACCTCAGGCACTTGGCATTTTCAACAGGA 1200
QY 1141 ----- 1140
Db 1201 CTGTCCGTATCTTCACACAGAGGCCCGCGAGACTCCCCCGCTGCCCTTACAC 1260
QY 1141 ----- ACCACTCCGGATTAATTTCTCAGCTGTACCCCATCAGCATTTTGGCCG 1191
Db 1261 CCTTCTACTACCCACTCCGGATTAATTTCTCAGCTGTACCCCATCAGCATTTTGGCCG 1320
QY 1192 TTCCGATCATCATCTTATTCAGAGCAGAGATTTGATCTTCAATCCCAACCAAG 1251
Db 1321 TTCCGATCATCATCTTATTCAGAGCAGAGATTTGATCTTCAATCCCAACCAAG 1380
QY 1252 GCTGTGGGCGCATCATCGGGAAGAAAGGGGCAACATCAAAAGCTGTGGAGATTGGC 1311
Db 1381 GCTGTGGGCGCATCATCGGGAAGAAAGGGGCAACATCAAAAGCTGTGGAGATTGGC 1440
QY 1312 GGAAGCTCTATCAAGATTGCCCCGCGAAAGGCCAGAGCTCAGCGAAAGATGTATC 1371
Db 1441 GGAAGCTCTATCAAGATTGCCCCGCGAAAGGCCAGAGCTCAGCGAAAGATGTATC 1500
QY 1372 ATCACCGGGCCACCGGAAACCCAGTTCAAGGCCAGGACAGATCTTTGGAAATCTGAAA 1431
Db 1501 ATCACCGGGCCACCGGAAACCCAGTTCAAGGCCAGGACAGATCTTTGGAAATCTGAAA 1560
QY 1432 GAGAAATCTTCTTTAACCCCAAGAAAGATGAACTGGAACGGATTAACAGAGGCC 1491
Db 1561 GAGAAATCTTCTTTAACCCCAAGAAAGATGAACTGGAACGGATTAACAGAGGCC 1620
QY 1492 TCTTCAACAGCTGGCGGGGATTTGGCAAGGTGGCAAGACCTGTGAACGATCTGAGAAC 1551
Db 1621 TCTTCAACAGCTGGCGGGGATTTGGCAAGGTGGCAAGACCTGTGAACGATCTGAGAAC 1680
QY 1552 TTAAACAGTGCAGAAATCATCTGCTGTGACCAAGCGCAATGAAATGAGAAAGTG 1611
Db 1681 TTAAACAGTGCAGAAATCATCTGCTGTGACCAAGCGCAATGAAATGAGAAAGTG 1740
QY 1612 ATGTCAAGATTAATGGGCACTTCTTTGCAAGCAATCTGCAAGGCAAGATCAAGGAA 1671
Db 1741 ATGTCAAGATTAATGGGCACTTCTTTGCAAGCAATCTGCAAGGCAAGATCAAGGAA 1800
QY 1672 ATTGTCAACAGTGAAGCAGAGCAGAGAGAGAAATACCTCAGGAGTGGCTCAAGCGC 1731
Db 1801 ATTGTCAACAGTGAAGCAGAGCAGAGAGAGAAATACCTCAGGAGTGGCTCAAGCGC 1860
QY 1732 AGCAAGTGAAGCTCCACAGGCAACAGCAAAACAAAGATGATGTAGCTTCCAAAC 1791
Db 1861 AGCAAGTGAAGCTCCACAGGCAACAGCAAAACAAAGATGATGTAGCTTCCAAAC 1920
QY 1792 CTGACAGATGAACCAAGCAGCAGCAGCAGATCCGGAACCAAGACCATCTGAG 1851
Db 1921 CTGACAGATGAACCAAGCAGCAGCAGCAGATCCGGAACCAAGACCATCTGAG 1980
QY 1852 GAATGAAGTCTGCGGAGCGGCAAGGACTCTGCGAGGCTCTGAAACCCAGGGGC 1911
Db 1981 GAATGAAGTCTGCGGAGCGGCAAGGACTCTGCGAGGCTCTGAAACCCAGGGGC 2040
QY 1912 CGAGAGAGGCGGGAAGGTACAGGCTTTGCAAGAACCAAGGCGCGCTCCCGCC 1971
Db 2041 CGAGAGAGGCGGGAAGGTACAGGCTTTGCAAGAACCAAGGCGCGCTCCCGCC 2100
QY 1972 CCCCAGGCTTTCTGAGGCTTCAAGCCATCACTTCAACATCACTGGATCTCTGAA 2031
Db 2101 CCCCAGGCTTTCTGAGGCTTCAAGCCATCACTTCAACATCACTGGATCTCTGAA 2160
QY 2032 CTCCACAGACGCTATCTCTTTAGTTGAATCAATAGGTGAACGTGTTCAAGCCAAGC 2091
Db 2161 CTCCACAGACGCTATCTCTTTAGTTGAATCAATAGGTGAACGTGTTCAAGCCAAGC 2220

QY 2092 AAAATGACACCCCTTTTCTGTGGCAATGCTCTGTACATGTGTATCATATTAGAAA 2151
|
|
|
Db 2221 AAAATGACACCCCTTTTCTGTGGCAATGCTCTGTACATGTGTATCATATTAGAAA 2280
|
|
|
QY 2152 GGGAGATGTAGATATGTGGCTGTGGTTACACAGGGTGCCTGCAGCGGTATATAT 2211
|
|
|
Db 2281 GGGAGATGTAGATATGTGGCTGTGGTTACACAGGGTGCCTGCAGCGGTATATAT 2340
|
|
|
QY 2212 TTAGAAATATATATCAATTAATCACTTAATCACTTAATTAATTAATTTT 2271
|
|
|
Db 2341 TTAGAAATATATATCAATTAATCACTTAATCACTTAATTAATTAATTTT 2400
|
|
|
QY 2272 TTTTCTTTTAAAGAAAGCAGGCTTTCTAGACTTTAAAGATTAAGCTTTGGAG 2331
|
|
|
Db 2401 TTTTCTTTTAAAGAAAGCAGGCTTTCTAGACTTTAAAGATTAAGCTTTGGAG 2460
|
|
|
QY 2332 GTCTCAGGCTGTAAGAGAGAGCTTTGAGGCCACCCGCACAAATTCACCCAGAGGAAAT 2391
|
|
|
Db 2461 GTCTCAGGCTGTAAGAGAGAGCTTTGAGGCCACCCGCACAAATTCACCCAGAGGAAAT 2520
|
|
|
QY 2392 CTCTCGAGAGGACACTCAGGCGAGTCTGATCACTGTATGTCAACGAAAGGATTA 2451
|
|
|
Db 2521 CTCTCGAGAGGACACTCAGGCGAGTCTGATCACTGTATGTCAACGAAAGGATTA 2580
|
|
|
QY 2452 CCGTCTCTTGAAGAGAACTGTGTCACTCTCAATGCTGTAGCTCATACACCCATT 2511
|
|
|
Db 2581 CCGTCTCTTGAAGAGAACTGTGTCACTCTCAATGCTGTAGCTCATACACCCATT 2640
|
|
|
QY 2512 TCTCTTTGCTTCAACAGTTTAAACGTGTTTGTGCATATGCTATATATTTCTGTCT 2571
|
|
|
Db 2641 TCTCTTTGCTTCAACAGTTTAAACGTGTTTGTGCATATGCTATATATTTCTGTCT 2700
|
|
|
QY 2572 CTCTCTGTTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2631
|
|
|
Db 2701 CTCTCTGTTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760
|
|
|
QY 2632 TTTCTCATATCCCTCATCTCAATCCGATCTAGACACCCCTCCCTCCCTCCCTCCCT 2691
|
|
|
Db 2761 TTTCTCATATCCCTCATCTCAATCCGATCTAGACACCCCTCCCTCCCTCCCTCCCT 2820
|
|
|
QY 2692 GTGCTCGAGTATCAATCAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2751
|
|
|
Db 2821 GTGCTCGAGTATCAATCAACAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 2880
|
|
|
QY 2752 CTACACCTGGTACTCAAAAGAAAGAAAGTCAATGATCTGTCTAGAGCTTTTGGAG 2811
|
|
|
Db 2881 CTACACCTGGTACTCAAAAGAAAGAAAGTCAATGATCTGTCTAGAGCTTTTGGAG 2940
|
|
|
QY 2812 AGGAAACAGGAACCCACCAACCAACCAATCAACCAACCAAGAAAAATTCACATG 2871
|
|
|
Db 2941 AGGAAACAGGAACCCACCAACCAACCAATCAACCAACCAAGAAAAATTCACATG 3000
|
|
|
QY 2872 AAAGAAAGTATTTTGTCTTTTGTGATTTGTGTATTAAGCCATATATTCAGAAATG 2931
|
|
|
Db 3001 AAAGAAAGTATTTTGTCTTTTGTGATTTGTGTATTAAGCCATATATTCAGAAATG 3060
|
|
|
QY 2932 ATTCTCTTCTTAAAAAATGAGAGAAAGTGAATTTTACCAAGTTTGGCC 2991
|
|
|
Db 3061 ATTCTCTTCTTAAAAAATGAGAGAAAGTGAATTTTACCAAGTTTGGCC 3120
|
|
|
QY 2992 CAGGGCGTTAATTCACAGATTTTAAAGAAAAACACAGAGAAAGCTACCTCAG 3051
|
|
|
Db 3121 CAGGGCGTTAATTCACAGATTTTAAAGAAAAACACAGAGAAAGCTACCTCAG 3180
|
|
|
QY 3052 GTGTTTAACTCAGACACCTGTCTGTGTCTTCCCTTAAGATTTTGAAGCTGATAG 3111
|
|
|
Db 3181 GTGTTTAACTCAGACACCTGTCTGTGTCTTCCCTTAAGATTTTGAAGCTGATAG 3240
|
|
|
QY 3112 TTGAGACATTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3171
|
|
|
Db 3241 TTGAGACATTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300
|
|
|
QY 3172 GCCAGCCTGAGAGAGGTGACAGTCCAGTGTGCAACAGCTGTGTAATGTCTCCGCT 3231
|
|
|

Db 3301 GCCAGCCTGAGAGAGGTGACAGTCCAGTGTGCAACAGCTGTGTAATGTCTCCGCT 3360
|
|
|
QY 3232 AGCCAAAGACCNATATGACCTTTTGTGACAAACCTTGAATGTTATTT 3283
|
|
|
Db 3361 AGCCAAAGACCNATATGACCTTTTGTGACAAACCTTGAATGTTATTT 3412
|
|
|
RESULT 5
AAS26150
ID AAS26150 standard, cDNA, 3694 BP.
XX
AC AAS26150;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 329.
KW
KW Human; immunosuppressive; antiarthritic; ss; antineumatic; cyostatic;
KW cardiatic; vasotropic; cerebroprotective; neurotropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnery;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cerebrovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN W020015322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205151P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.

QY 128 GGCAGCTTTTGGGACAGAAAGCTGCCCTGGCGGGACAGGTCTCTGAAGTCCGCT 187
Db 151 GGCAGCTTTTGGGACAGAAAGCTGCCCTGGCGGGACAGGTCTCTGAAGTCCGCT 210
QY 188 AGGCTTTCGTGACCTACCCGACAGAACTGGGCACTCCGCGCATGAGACCTCTCGG 247
Db 211 AGGCTTTCGTGACCTACCCGACAGAACTGGGCACTCCGCGCATGAGACCTCTCGG 270
QY 248 GTAAAGTGAATTGACATGGGAAATCATGGAAGTTGATTACTCAGTCTTAAAAAGCTTA 307
Db 271 GTAAAGTGAATTGACATGGGAAATCATGGAAGTTGATTACTCAGTCTTAAAAAGCTTA 330
QY 308 GGAGGAGGAAAAATTCAGATTGAAACATCCCTCTCACTGCACTGGAGGGAGGTGGAAATG 367
Db 331 GGAGGAGGAAAAATTCAGATTGAAACATCCCTCTCACTGCACTGGAGGGAGGTGGAAATG 390
QY 368 GACTTTTGGCTCAATATGAGACAGTGAAGATGTGAAAATGTCACACAGACAGAA 427
Db 391 GACTTTTGGCTCAATATGAGACAGTGAAGATGTGAAAATGTCACACAGACAGAA 450
QY 428 CGGCGGTGTCAAGCTCATATGCAACAGAGAAAGCAAAATATGCAATGAGAAC 487
Db 451 CGGCGGTGTCAAGCTCATATGCAACAGAGAAAGCAAAATATGCAATGAGAAC 510
QY 488 TAAAGGCGCATCAGTTTGAAGACTACTCTTCAAGATTCTTACATCCCGATGAAGAG 547
Db 511 TAAAGGCGCATCAGTTTGAAGACTACTCTTCAAGATTCTTACATCCCGATGAAGAG 570
QY 548 TGAAGTCCCTTTCGCCCCCTTCAAGCGAGCCAGCGTGGGAGCACTCTTCCGCGAGCAAG 607
Db 571 TGAAGTCCCTTTCGCCCCCTTCAAGCGAGCCAGCGTGGGAGCACTCTTCCGCGAGCAAG 630
QY 608 GGCAGCGCCCTGGGGGGCACTTCTCAGGCGACACAGATTGATTTCGCGTGGGATCTGG 667
Db 631 GGCAGCGCCCTGGGGGGCACTTCTCAGGCGACACAGATTGATTTCGCGTGGGATCTGG 690
QY 668 TCCCCACCAGTTTGTGTGGTGCATCATCGGAAAGAGGGCTTGAACATAAAGAACATCA 727
Db 691 TCCCCACCAGTTTGTGTGGTGCATCATCGGAAAGAGGGCTTGAACATAAAGAACATCA 750
QY 728 CTAAAGCAGACCCAGTCCCGGGTAGATATCATAGAAAAGAACTCTGGAAGCTGACAGA 787
Db 751 CTAAAGCAGACCCAGTCCCGGGTAGATATCATAGAAAAGAACTCTGGAAGCTGACAGA 810
QY 788 AGCGTGCACATCCATGCGCACCCGAGGGGACTTCTGAAGCATGCCGATGATTCTTG 847
Db 811 AGCGTGCACATCCATGCGCACCCGAGGGGACTTCTGAAGCATGCCGATGATTCTTG 870
QY 848 AATCATGCAAGAAAGGAGCAGATGAGCAAACTAGCCGAGAGATTCTCTGAAAAATCT 907
Db 871 AATCATGCAAGAAAGGAGCAGATGAGCAAACTAGCCGAGAGATTCTCTGAAAAATCT 930
QY 908 TGGCACAATGCTGTGTGGTGAAGACTGATTGAAAAAGAGGACAAATTTGAAGAAA 967
Db 931 TGGCACAATGCTGTGTGGTGAAGACTGATTGAAAAAGAGGACAAATTTGAAGAAA 990
QY 968 TTGAACATGAAACAGGGACCAAGTAACTCATCTTTGCAAGATTTGCAATATCA 1027
Db 991 TTGAACATGAAACAGGGACCAAGTAACTCATCTTTGCAAGATTTGCAATATCA 1050
QY 1028 ACCGGAAGAAACATCAGTGTGAAGGACAGTTGAGGCTTGAGCAAGTCTGAGATAG 1087
Db 1051 ACCGGAAGAAACATCAGTGTGAAGGACAGTTGAGGCTTGAGCAAGTCTGAGATAG 1110
QY 1088 AGATTATGAAGAGCTGCTGAAGGCTTTTGAATAATGATGCTGCTGTAAAC----- 1140
Db 1111 AGATTATGAAGAGCTGCTGAAGGCTTTTGAATAATGATGCTGCTGTAAACAGAG 1170
QY 1141 ----- 1140
Db 1171 CCAATCTGATCCAGGCTTGAACCTCAGCGCAGCTTGGCATCTTTCAACAGAGCTGTCCG 1230
QY 1141 ----- 1140

Db 1231 TGTATATCTCACACAGAGGCCCCGGAGACTCCCCCGGCTGCCCCCTACACACCCCTTCA 1290
QY 1141 --ACCCACTCCGGATACTTCTTCAGCCTGTACCCCCATCACAGATTGGCCGCTCCGCG 1198
Db 1291 CTACCCACTCCGGATACTTCTTCAGCCTGTACCCCCATCACAGATTGGCCGCTCCGCG 1350
QY 1199 ATCATCATCTTTATCAAGACAGAGATTTGATATCTTCTTCAATCCCAACCAAGCTGTGG 1258
Db 1351 ATCATCATCTTTATCAAGACAGAGATTTGATATCTTCTTCAATCCCAACCAAGCTGTGG 1410
QY 1259 GGGCATATGAGGAGAAAGGGGGGACACATCAAACTGCTGGCGAGATTTCCGCGAGCCT 1318
Db 1411 GGGCATATGAGGAGAAAGGGGGGACACATCAAACTGCTGGCGAGATTTCCGCGAGCCT 1470
QY 1319 CTATCAAGATTGCCCTCTCGAAGGCGCCAGACGTGAGGAAAGATGTCATCATCACCG 1378
Db 1471 CTATCAAGATTGCCCTCTCGAAGGCGCCAGACGTGAGGAAAGATGTCATCATCACCG 1530
QY 1379 GGCCACCGGAAGCCCATGTTCAAGGCTCAGGAGCCGATCTTTGGGAACTGAAAGAGAAA 1438
Db 1531 GGCCACCGGAAGCCCATGTTCAAGGCGCCAGGACCGATCTTTGGGAACTGAAAGAGAAA 1590
QY 1439 ACTCTTTAACCCCAAGAGAGATGAGAGCTGGAAGCGCATATGCAAGTCCCTCTTCCA 1498
Db 1591 ACTCTTTAACCCCAAGAGAGATGAGAGCTGGAAGCGCATATGCAAGTCCCTCTTCCA 1650
QY 1499 CAGCTGGCCGGGTGATTTGGCAAGGTGGCAAGACCGTGAAGAACTGACAGAACTTTACA 1558
Db 1651 CAGCTGGCCGGGTGATTTGGCAAGGTGGCAAGACCGTGAAGAACTGACAGAACTTTACA 1710
QY 1559 GTGCAAGAAATCATGCTGTGCTGTGACCAAAACCGCAGATGAATAATGAGAGATGCTCA 1618
Db 1711 GTGCAAGAAATCATGCTGTGCTGTGACCAAAACCGCAGATGAATAATGAGAGATGCTCA 1770
QY 1619 GAATTTATCGGACCTTCTTGTGCGCAGATGCAAGCGGCAAGATGAGGAAATTTGAC 1678
Db 1771 GAATTTATCGGACCTTCTTGTGCGCAGATGCAAGCGGCAAGATGAGGAAATTTGAC 1830
QY 1679 AACAGTGAACACAGAGACAGAAATACCTTCAGGAGTGTGCTTCACAGCGCAGCAAGT 1738
Db 1831 AACAGTGAACACAGAGACAGAAATACCTTCAGGAGTGTGCTTCACAGCGCAGCAAGT 1890
QY 1739 GAGGCTCCACAGGACCAAGCAAAACAAAGGATATGTAGCCTTTCACACCTGACAG 1798
Db 1891 GAGGCTCCACAGGACCAAGCAAAACAAAGGATATGTAGCCTTTCACACCTGACAG 1950
QY 1799 AATGAGACAAAGCAGACAGCCGATCCGGAAGCAAAACCAAGACATCTTGAAGATGAG 1858
Db 1951 AATGAGACAAAGCAGACAGCCGATCCGGAAGCAAAACCAAGACATCTTGAAGATGAG 2010
QY 1859 AAGTCTCGAGAGGCGGACAGGACTTGCAGAGGCTTGAGAAACCCCAAGGGGCGGAGAG 1918
Db 2011 AAGTCTCGAGAGGCGGACAGGACTTGCAGAGGCTTGAGAAACCCCAAGGGGCGGAGAG 2070
QY 1919 GGGCGGGAAAGGTGACGAGGTTTGCAGAAACACAGAGCCCGGCTTCCGCGCCCAAG 1978
Db 2071 GGGCGGGAAAGGTGACGAGGTTTGCAGAAACACAGAGCCCGGCTTCCGCGCCCAAG 2130
QY 1979 GCTTCTGACAGGCTTCAAGCATCACTTCAACATCCATCCGATCTCTCTGAATCTCCAC 2038
Db 2131 GCTTCTGACAGGCTTCAAGCATCACTTCAACATCCATCCGATCTCTCTGAATCTCCAC 2190
QY 2039 GACGCTATCCCTTTTATTTGAACCTAATAGAGTGAAGCTGTTCAAAACCAAGCAAAATGC 2098
Db 2191 GACGCTATCCCTTTTATTTGAACCTAATAGAGTGAAGCTGTTCAAAACCAAGCAAAATGC 2250
QY 2099 ACACCTTTTCTGTGCAAAATGCTCTGTGATCTGTGTATCATATTATGAAGGAGAGA 2158
Db 2251 ACACCTTTTCTGTGCAAAATGCTCTGTGATCTGTGTATCATATTATGAAGGAGAGA 2310
QY 2159 TGTTAAGATATGAGGCTGTGGGTTTACAGGGGCTGTGACGGGTAAATATTTTGA 2218

PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251856P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
PI Rosen CA, Ruben SM, Barash SC;
XX WPI: 2003-147444/14.
DR P-PSDB; ABUS5231.
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
FT renal disorders.

XX Claim 1; SEQ ID NO 329; 402bp; English.

XX The invention relates to human novel polypeptides and their associated
XX polynucleotides. The polypeptides and polynucleotides are useful in gene
XX therapy for treating, inhibiting or preventing neural disorders, immune
XX system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
XX and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
XX nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
XX gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
XX (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
XX heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
XX renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
XX leukemia), inflammatory diseases (e.g. septic shock, bursitis and
XX appendicitis), allergic reactions and conditions (e.g. asthma), blood
XX related disorders (e.g. thrombosis, atherosclerosis and myocardial
XX infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
XX human novel polynucleotides of the invention

XX Sequence 3694 BD; 1103 A; 934 C; 839 G; 817 T; 0 U; 1 Other;

XX Query Match 93.0%; Score 3052.6; DB 7; Length 3694;
XX Best Local Similarity 96.0%; Pred. No. 0;

XX Matches 3214; Conservative 1; Mismatches 1; Indels 132; Gaps 3;

QY 68 GGATGATGAACAAGCTTTAATCGGGAAGCTGAGCCCGCCGCTACACCGCGAGCACTCC 127
DB 91 GGATGATGAACAAGCTTTAATCGGGAAGCTGAGCCCGCCGCTACACCGCGAGCACTCC 150
QY 128 GGGAGCTCTTTGGGAGCAGGAAGCTGCCCTGGCGGAGCAGTCTCTGTAAGTCCGGCT 187
DB 151 GGGAGCTCTTTGGGAGCAGGAAGCTGCCCTGGCGGAGCAGTCTCTGTAAGTCCGGCT 210
QY 188 ACCGCTTCCTGAGCACTACCCGAGCAGGAAGTGGCGCATCCGCCGATGAGAACCTCTCGG 247
DB 211 ACCGCTTCCTGAGCACTACCCGAGCAGGAAGTGGCGCATCCGCCGATGAGAACCTCTCGG 270
QY 248 GTAAAGTGAATTCATGAGGGAATATGAGAGTGAATTAATCAAGTCTTAAAAAGTTAA 307
DB 271 GTAAAGTGAATTCATGAGGGAATATGAGAGTGAATTAATCAAGTCTTAAAAAGTTAA 330
QY 308 GGAGCAGAAAAATTCAGATTCGAAACATCCCTCTCACTGACGTGGAGGATGTTGGATG 367
DB 331 GGAGCAGAAAAATTCAGATTCGAAACATCCCTCTCACTGACGTGGAGGATGTTGGATG 390
QY 368 GATCTTTGGCTCAATATGGGAGCAGTGAAGATGTGAACAAGTCAACACAGACAGAAA 427
DB 391 GATCTTTGGCTCAATATGGGAGCAGTGAAGATGTGAACAAGTCAACACAGACAGAAA 450
QY 428 CGCCGCTTGTCAAGTCACATATGCAACAAGAGAAAGAAATATGCAATGAGAGAGC 487
DB 451 CGCCGCTTGTCAAGTCACATATGCAACAAGAGAAAGAAATATGCAATGAGAGAGC 510
QY 488 TAAAGCGGAGCATAGTTGAGAACTACTCTTCAAGATTTCTCAATCCGAGTGAAGAG 547

DB 511 TAAAGCGGAGCATAGTTGAGAACTACTCTTCAAGATTTCTCAATCCGAGTGAAGAG 570
QY 548 TGAAGTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGAGCACTCTTCCGGGAGCAAG 607
DB 571 TGAAGTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGAGCACTCTTCCGGGAGCAAG 630
QY 608 GCGACGCCCCCTGGGGGCACTTCTCAGGCGCAACAGATGATTTCCCGCTGGGATCTCGG 667
DB 631 GCGACGCCCCCTGGGGGCACTTCTCAGGCGCAACAGATGATTTCCCGCTGGGATCTCGG 690
QY 668 TCCCAACCAATTTGTTGGTGCATCATCGGAAAGGAGGCTTGACCATTAAGAACATCA 727
DB 691 TCCCAACCAATTTGTTGGTGCATCATCGGAAAGGAGGCTTGACCATTAAGAACATCA 750
QY 728 CTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAAGAAACTCTGAGCTGACAGAG 787
DB 751 CTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAAGAAACTCTGAGCTGACAGAG 810
QY 788 AGCTGTACCATCATGCCACCCCAAGGGGAGCTTTGAAAGCATGCCGATGATTCCTG 847
DB 811 AGCTGTACCATCATGCCACCCCAAGGGGAGCTTTGAAAGCATGCCGATGATTCCTG 870
QY 848 AATCATGCAAGAAAGAGGAGATGAGACCAACTAGCGGAGAGATTCCTGAAAAATCT 907
DB 871 AATCATGCAAGAAAGAGGAGATGAGACCAACTAGCGGAGAGATTCCTGAAAAATCT 930
QY 908 TGGCACAATAGCTTGGTGAAGAGTGAATTTGAAAAAGAGGAGAAATTTGAAGAAA 967
DB 931 TGGCACAATAGCTTGGTGAAGAGTGAATTTGAAAAAGAGGAGAAATTTGAAGAAA 990
QY 968 TTGAACATGAAACAGGAGCAAGATTAACAATCTCATCTTTGACAGATTTGAGATTAACA 1027
DB 991 TTGAACATGAAACAGGAGCAAGATTAACAATCTCATCTTTGACAGATTTGAGATTAACA 1050
QY 1028 ACCCGGAAAGACATCATCTGTGAAGGCGACAGTTGAGGCTGTGCCAGTCTGAGATAG 1087
DB 1051 ACCCGGAAAGACATCATCTGTGAAGGCGACAGTTGAGGCTGTGCCAGTCTGAGATAG 1110
QY 1088 AGATTATGAAGAGCTGCGTAGAGCTTTGAAAAATGATAGCTGGCTTTAAC----- 1140
DB 1111 AGATTATGAAGAGCTGCGTAGAGCTTTGAAAAATGATAGCTGGCTTTAACCAAGAG 1170
QY 1141 ----- 1140
DB 1171 CCAATGTATCCAGGGTGAACCTCAGCGCACTTGAGCATCTTTCAACAGAGCTGCG 1230
QY 1141 ----- 1140
DB 1231 TGCATATCCACAGAGAGGCCCCGGAGAGCTCCCCCGCTGCCCCCTACACCCCTTCA 1290
QY 1141 --ACCACTCCGAGATCTTCCAGGCTGTACCCCATCAACAGTTGAGCCGTTCCGCG 1198
DB 1291 CTACCCAGCTCGGATATCTTCCAGGCTGTACCCCATCAACAGTTGAGCCGTTCCGCG 1350
QY 1199 ATCATCACTTTATCCAGAGCAGAGATGTGAATCTTTCAATCCAAACCAAGCTGTGG 1258
DB 1351 ATCATCACTTTATCCAGAGCAGAGATGTGAATCTTTCAATCCAAACCAAGCTGTGG 1410
QY 1259 GGGCATTATCCGGAAGAGAGGGGGCACAATCAAAAGCTGGGAGAAATTTGCGGAGCTT 1318
DB 1411 GGGCATTATCCGGAAGAGAGGGGGCACAATCAAAAGCTGGGAGAAATTTGCGGAGCTT 1470
QY 1319 CTATCAAGATTGCCCCCTGGGAGAGGCCAGAGCTCAGCGAAAGATGTCATCATCACCG 1378
DB 1471 CTATCAAGATTGCCCCCTGGGAGAGGCCAGAGCTCAGCGAAAGATGTCATCATCACCG 1530
QY 1379 GGGCACCAGAGAGCTCGATTCAAGGCCCAAGGAGCGATCTTTGGGAAATTTGAAGAGAAA 1438
DB 1531 GGGCACCAGAGAGCTCGATTCAAGGCCCAAGGAGCGATCTTTGGGAAATTTGAAGAGAAA 1590
QY 1439 ACTTCTTTAACCCCAAGAGAGAGTGAAGCTGGAGCGGATATCAAGTGGCCCTCTTCCA 1498

XX 02-AUG-2001.
PD 17-JAN-2001; 2001WO-US001341.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-018974P.
PR 17-MAR-2000; 2000US-019076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 03-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233397P.
PR 14-SEP-2000; 2000US-0233398P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241828P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244618P.
PR 08-NOV-2000; 2000US-0244619P.
PR 08-NOV-2000; 2000US-0244620P.
PR 08-NOV-2000; 2000US-0244621P.
PR 08-NOV-2000; 2000US-0244622P.
PR 08-NOV-2000; 2000US-0244623P.
PR 08-NOV-2000; 2000US-0244624P.
PR 08-NOV-2000; 2000US-0244625P.
PR 08-NOV-2000; 2000US-0244626P.
PR 08-NOV-2000; 2000US-0244627P.
PR 08-NOV-2000; 2000US-0244628P.
PR 08-NOV-2000; 2000US-0244629P.
PR 08-NOV-2000; 2000US-0244630P.
PR 08-NOV-2000; 2000US-0244631P.
PR 08-NOV-2000; 2000US-0244632P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251889P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Barash SC, Ruben SM;
PI XX

DR WPI: 2001-48783/53.
DR P-PSDB; AAIU6166.

PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.

PS Claim 1; SEQ ID NO 332; 980bp; English.

XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. Rheumatoid arthritis, hyperproliferative disorders e.g. cardiac
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 60.6%; Score 1989.8; DB 4; Length 2290;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2028; Conservative 0; Mismatches 8; Indels 4; Gaps 3;

QY	1247	CCCAAGCTGTGGGCGCATCATCGGGAAGAAGGGGGGACATCAACAGCTGGCGAGAT	1306
DB	25	CCCAAGCTGTGGGCGCATCATCGGGAAGAAGGGGGGACATCAACAGCTGGCGAGAT	84
QY	1307	TCGCGGAGGCTTATCAAGATTGCGCTGCGAAGGCCCAAGCTGACGGAAAGATGG	1366
DB	85	TCGCGGAGGCTTATCAAGATTGCGCTGCGAAGGCCCAAGCTGACGGAAAGATGG	144
QY	1367	TCATCATCACCGGGCCACCGGAGCCCAAGTTCAAGGCCCGGAGCGATTTTGGGAAC	1426
DB	145	TCATCATCACCGGGCCACCGGAGCCCAAGTTCAAGGCCCGGAGCGATTTTGGGAAC	204
QY	1427	TGAAGAAGAAAACTTCTTAACCCCAAGAAAGTGAAGTGAAGCGCATATACAG	1486
DB	205	TGAAGAAGAAAACTTCTTAACCCCAAGAAAGTGAAGTGAAGCGCATATACAG	264
QY	1487	TGCCCTTTCCACAGCTGCGCGGGTGAATGCAAAAGGTGCAAGACCTGAAAGACTGC	1546
DB	265	TGCCCTTTCCACAGCTGCGCGGGTGAATGCAAAAGGTGCAAGACCTGAAAGACTGC	324
QY	1547	AGAACTTAACCAAGTGAAGTCACTGCTGCTGCAAAAGGCCAGATGAAGAAATGAG	1606
DB	325	AGAACTTAACCAAGTGAAGTCACTGCTGCTGCAAAAGGCCAGATGAAGAAATGAG	384
QY	1607	AAGTGAATGTCAGAAATATGCGGCACTTTCTTGTACCGAGCTGCAAGCGCAAGTCA	1666
DB	385	AAGTGAATGTCAGAAATATGCGGCACTTTCTTGTACCGAGCTGCAAGCGCAAGTCA	444
QY	1667	GGAAGATTTGTAACAAGGTGAAGCAAGCAAGAAATACCTTCAGGAGTGCCTTCAC	1726
DB	445	GGAAGATTTGTAACAAGGTGAAGCAAGCAAGAAATACCTTCAGGAGTGCCTTCAC	504
QY	1727	AGCGCAGCAAGTGAAGTCCCAACAGGCAACAGCAAAACAAACGATGATGAAGCCCTTC	1786
DB	505	AGCGCAGCAAGTGAAGTCCCAACAGGCAACAGCAAAACAAACGATGATGAAGCCCTTC	564

QY	1787	AACACCTGACAGATGAGACCAAAAGCGACCCAGCCAGATCGGAGGCAAAACAAAGCCAT	1846
DB	565	AACACCTGACAGATGAGACCAAAAGCGACCCAGCCAGATCGGAGGCAAAACAAAGCCAT	624
QY	1847	CTGAGGAATGAAGAAGTCTCGGAGGCGGCGGAGGAGTCTCGCGAAGGCCCTGGAACCCCA	1906
DB	625	CTGAGGAATGAAGAAGTCTCGGAGGCGGCGGAGGAGTCTCGCGAAGGCCCTGGAACCCCA	684
QY	1907	GGGAGCCGAGAGGGGGGGGAGAGTCAAGCAGATTGTCAGAAACCAACCGAGCCCGCTTC	1966
DB	685	GGGAGCCGAGAGGGGGGGGAGAGTCAAGCAGATTGTCAGAAACCAACCGAGCCCGCTTC	744
QY	1967	CCGCCCCCGAGGGCTTTGCAAGGCTTCAAGCCATCACTTCACTCACTGAGATCTTC	2026
DB	745	CCGCCCCCGAGGGCTTTGCAAGGCTTCAAGCCATCACTTCACTCACTGAGATCTTC	804
QY	2027	CTGAATCTCCACGAGCTATCCCTTTTAACTGAACATATAGTGAAGGTTCAAGC	2086
DB	805	CTGAATCTCCACGAGCTATCCCTTTTAACTGAACATATAGTGAAGGTTCAAGC	864
QY	2087	CAAGCAAAATGACACACCTTTTCTGTGGCAAAATGCTCTGTATCAATGTGTATATAT	2146
DB	865	CAAGCAAAATGACACACCTTTTCTGTGGCAAAATGCTCTGTATCAATGTGTATATAT	924
QY	2147	AGAAAGGAAAGATGTTAAGATATGTGGCTGTGGTTTACACAGGCTGCGAGGTTAA	2206
DB	925	AGAAAGGAAAGATGTTAAGATATGTGGCTGTGGTTTACACAGGCTGCGAGGTTAA	984
QY	2207	TATATTTGAAAT	2266
DB	985	TATATTTGAAAT	1044
QY	2267	ATTTTCTTTCTTTTAAAGAGAAAGAGGCTTTTCTGAGCTTTTAAAGTAAAGTCTTT	2326
DB	1045	ATTTTCTTTCTTTTAAAGAGAAAGAGGCTTTTCTGAGCTTTTAAAGTAAAGTCTTT	1104
QY	2327	GGGAGTCTCAAGGTGTAGAGAGAGCTTTGAGGCCACCGGCAAAATTCACCCAGAG	2386
DB	1105	GGGAGTCTCAAGGTGTAGAGAGAGCTTTGAGGCCACCGGCAAAATTCACCCAGAG	1164
QY	2387	GAAATCTCGTGGAGAGACACACGAGGAGTGTGATCACTGTGTATGTCAACAGAG	2446
DB	1165	GAAATCTCGTGGAGAGACACACGAGGAGTGTGATCACTGTGTATGTCAACAGAG	1224
QY	2447	GGATACCGTCTCTTGAAGAGAGAACTGTGTCACTCTCATGCTGTAGTCAATAC	2506
DB	1225	GGATACCGTCTCTTGAAGAGAGAACTGTGTCACTCTCATGCTGTAGTCAATAC	1284
QY	2507	CAATTTCTTTGCTTCAAGGTTTAACTGTTTGTGATATGCTATATATATCTC	2566
DB	1285	CAATTTCTTTGCTTCAAGGTTTAACTGTTTGTGATATGCTATATATATCTC	1344
QY	2567	TGTCTCTCTGTATATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCT	2626
DB	1345	TGTCTCTCTGTATATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCT	1403
QY	2627	TGGAATTTCTTCAATCCCTTCATCTCAATCCCGATCTAGCAAC - CCCCCCCCCCAGG	2684
DB	1404	TGGAATTTCTTCAATCCCTTCATCTCAATCCCGATCTAGCAAC CCCCCCCCCCAGG	1463
QY	2685	CAAAAGCTGCTGAGTATCAATCAACAAAGAAACAAAGGAAACAAACCAACCA	2744
DB	1464	CAAAAGCTGCTGAGTATCAATCAACAAAGAAACAAAGGAAACAAACCAACCA	1523
QY	2745	GCTCAACTTCACTGGTACTCAAAAGAAACAGAGTCAATGGTACTTCTAGAGCTT	2804
DB	1524	GCTCAACTTCACTGGTACTCAAAAGAAACAGAGTCAATGGTACTTCTAGAGCTT	1583
QY	2805	TTGGAAGAGAGAAACAGAAACCAACCAACCAATCAACCAACCAAGAAATTC	2864
DB	1584	TTGGAAGAGAGAAACAGAAACCAACCAACCAATCAACCAACCAAGAAATTC	1643

QY 2865 CACAAATGAAAGATGATATTTGCTTTTGGATTTGGTATTAAGCCATCATATATTCAG 2924
DB 1644 CACAAATGAAAGATGATATTTGCTTTTGGATTTGGTATTAAGCCATCATATATTCAG 1703
QY 2925 CAAATGATCTCTTTCTTT-AAAAAATAATGAGAGAAAGTAAATTTACCAAGT 2983
DB 1704 CAAATGATCTCTTTCTTTAAAAAATAATGAGAGAAAGTAAATTTACCAAGT 1763
QY 2984 TGTGGCCAGGAGGCTTAAATTCAGAGATTTTAAACGAGAAAAACACAGAAAGAC 3043
DB 1764 TGTGGCCAGGAGGCTTAAATTCAGAGATTTTAAACGAGAAAAACACAGAAAGAC 1823
QY 3044 TACCTCAGGTGTTTTTACCTCAGACCTTGCTGTTGTTCCCTTAGAGATTTGTAA 3103
DB 1824 TACCTCAGGTGTTTTTACCTCAGACCTTGCTGTTGTTCCCTTAGAGATTTGTAA 1883
QY 3104 GCTGATGATGAGACATTTTATTTTATTTATTAATAAATGATGAAAAAATAAGA 3163
DB 1884 GCTGATGATGAGACATTTTATTTTATTTATTAATAAATGATGAAAAAATAAGA 1943
QY 3164 TATCAACTGCAGCCTGAGAGAGTGAAGTCCAAAGTGAACAGCTGTTCTGAATTGT 3223
DB 1944 TATCAACTGCAGCCTGAGAGAGTGAAGTGAACAGCTGTTCTGAATTGT 2003
QY 3224 CTTCGCTAGCCAGAAACATATATGCTCTTTTGGACAAACCTTGAAAAATGTTATTT 3283
DB 2004 CTTCGCTAGCCAGAAACATATATGCTCTTTTGGACAAACCTTGAAAAATGTTATTT 2063

RESULT 8
ABX73494
ID ABX73494 standard; DNA; 2290 BP.
AC ABX73494;
XX
XX
XX 18-MAR-2003 (first entry)
XX
XX
XX Human novel polynucleotide #322.
DE
XX
XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antinflammatory;
KW cardiovascular; nephrotoxic; cytostatic; antiallergic; chromolytic;
KW haemostatic; antiatherosclerotic.
OS Homo sapiens.
XX
XX
XX US2002132753-A1.
XX
XX
XX 19-SEP-2002.
XX
XX
XX 17-JAN-2001; 2001US-00764864.
XX
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-021680P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225747P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.

PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249239P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
PA (RUBEN/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-14744/14.
DR P-PSDB; ABUS5234.
XX
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
XX
XX Claim 1; SEQ ID NO 332; 402pp; English.
XX
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
XX
SQ Sequence 2290 BP; 707 A; 561 C; 491 G; 531 T; 0 U; 0 Other;

Query Match 60.6%; Score 1989.8; DB 7; Length 2290;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2028; Conservative 0; Mismatches 8; Indels 4; Gaps 3;
QY 1247 CCAGGCTGTGGGCGCATCATCGGAGAGAGGGGCGCACATCAACAGCTGGCAGAT 1306
|||||

Db 25 CCGACGCTCCGGGCCATCATCGGAGAGAGGGGCAACATCAACAGCTGGCAGAT 84
QY 1307 TCGCCGAGGCTCTATCAAGATTGCCCCCTGCGGAGGCCCAAGCGTCAAGCCGAAAGATGG 1366
Db 85 TCGCCGAGGCTCTATCAAGATTGCCCCCTGCGGAGGGCCAGACCTCAAGCCGAAAGATGG 144
QY 1367 TCATCATCAAGGGGCGACCGGAGGCCAGTTCAAGGCCAGGAGCGGATCTTTGGGAAAC 1426
Db 145 TCATCATCAAGGGGCGACCGGAGGCCAGTTCAAGGCCAGGAGCGGATCTTTGGGAAAC 204
QY 1427 TGAAGAGGAAAACTTTTAAACCCCAAGAGAAAGTGAAGCTGGAAGCGCATATCAAG 1486
Db 205 TGAAGAGGAAAACTTTTAAACCCCAAGAGAAAGTGAAGCTGGAAGCGCATATCAAG 264
QY 1487 TGCCTCTTCCACAGCTGGCCGGGTGATTGGCAAAAGTGGCAAGCCGGTGAAGCAATGCG 1546
Db 265 TGCCTCTTCCACAGCTGGCCGGGTGATTGGCAAAAGTGGCAAGCCGGTGAAGCAATGCG 324
QY 1547 AGAAGCTTAACAGTGAAGTCAATGCTGCTGTAACCAAGCCAGATGAAGTGAAG 1606
Db 325 AGAAGCTTAACAGTGAAGTCAATGCTGCTGTAACCAAGCCAGATGAAGTGAAG 384
QY 1607 AAGTATCGTCAGAAATTAATCGGCACTTCTTTGTAACCAAGTGAAGTGAAGTGAAG 1666
Db 385 AAGTATCGTCAGAAATTAATCGGCACTTCTTTGTAACCAAGTGAAGTGAAGTGAAG 444
QY 1667 GGGAAATTGTACAAGGTGAAGCAGCAGGAGCAAGAAATACCTCAAGGGAGTGGCTTCA 1726
Db 445 GGGAAATTGTACAAGGTGAAGCAGCAGGAGCAAGAAATACCTCAAGGGAGTGGCTTCA 504
QY 1727 ACGGAGCAAGTGAAGGCTCCCAAGGCAACAGCAAAACAGGATGAATGAAGTGAAG 1786
Db 505 ACGGAGCAAGTGAAGGCTCCCAAGGCAACAGCAAAACAGGATGAATGAAGTGAAG 564
QY 1787 AACACCTGAAGAAATGAACAAACAGCAGCCAGTCGAGTGAAGTGAAGTGAAGTGAAG 1846
Db 565 AACACCTGAAGAAATGAACAAACAGCAGCCAGTCGAGTGAAGTGAAGTGAAGTGAAG 624
QY 1847 CTGAGGAATGAAGTCTGCGAGGGGCGGAGGAGTCTGCGAGGCGCTGAAGTGAAG 1906
Db 625 CTGAGGAATGAAGTCTGCGAGGGGCGGAGGAGTCTGCGAGGCGCTGAAGTGAAG 684
QY 1907 GGGGCGGAGGAGGGGCGGAGGAGTCAAGCAGGTTTGCAGAAACCAAGGCCGCTC 1966
Db 685 GGGGCGGAGGAGGGGCGGAGGAGTCAAGCAGGTTTGCAGAAACCAAGGCCGCTC 744
QY 1967 CCGCCCCCAGGGCTTCTGAGGCTTCAAGCATCACTTCAACATCACTCGGATCTC 2026
Db 745 CCGCCCCCAGGGCTTCTGAGGCTTCAAGCATCACTTCAACATCACTCGGATCTC 804
QY 2027 CTGAATCTCCAGACGCTATCCCTTTTGAATTGAATCAATAGTGAAGTGAAGTGAAG 2086
Db 805 CTGAATCTCCAGACGCTATCCCTTTTGAATTGAATCAATAGTGAAGTGAAGTGAAG 864
QY 2087 CAAGCAAAATGACACACCTTTTCTGAGCAAAATCGTCTGTACATGTGTACATATT 2146
Db 865 CAAGCAAAATGACACACCTTTTCTGAGCAAAATCGTCTGTACATGTGTACATATT 924
QY 2147 AGAAGGGAAGATGTTAAGATATGTGCTGTGTGTTACACAGGTCCTCGACGCGTAA 2206
Db 925 AGAAGGGAAGATGTTAAGATATGTGCTGTGTGTTACACAGGTCCTCGACGCGTAA 984
QY 2207 TATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2266
Db 985 TATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1044
QY 2267 AATTTTCTTTTAAAGAGAGGCTTTCTAGATTGAATTAAGATTAAGATTTT 2326
Db 1045 AATTTTCTTTTAAAGAGAGGCTTTCTAGATTGAATTAAGATTTTAAAGATTTT 1104
QY 2327 GGGAGGTCTACAGGTGTAGAGAGGCTTTGAGGCAACCGGCAAAATTCACCAAGG 2386
Db 1105 GGGAGGTCTACAGGTGTAGAGAGGCTTTGAGGCAACCGGCAAAATTCACCAAGG 1164

QY 2387 GAAATCTGTGGAAGGACACTCAAGGAGTCTGATCACTGTGTATGTCAACAGAG 2446
Db 1165 GAAATCTGTGGAAGGACACTCAAGGAGTCTGATCACTGTGTATGTCAACAGAG 1224
QY 2447 GATATCGGTCTTTGAAGAGAAACTGTATCACTCCCAAGCCGTGTAGTCAATAC 2506
Db 1225 GATATCGGTCTTTGAAGAGAAACTGTATCACTCCCAAGCCGTGTAGTCAATAC 1284
QY 2507 CCATTTCTTTTGTCTCAAGGTTTAACTGATTTTGTGACTGTATTAATTTCTC 2566
Db 1285 CCATTTCTTTTGTCTCAAGGTTTAACTGATTTTGTGACTGTATTAATTTCTC 1344
QY 2567 TGTCTCTGTGTATATCTCTCCCTCCCTCCCTCCCTCTTCTCCATCTTCATTTCT 2626
Db 1345 TGTCTCTGTGTATATCTCTCCCTCCCTCCCTCCCTCTTCTCCATCTTCATTTCT 1403
QY 2627 TTGAATTTCTCATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 2684
Db 1404 TTGAATTTCTCATCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1463
QY 2685 CAAAGCAGTCTGTAGTATCATATCAACAAAGGAAACAAAGCGAAACACAAACCA 2744
Db 1464 CAAAGCAGTCTGTAGTATCATATCAACAAAGGAAACAAAGCGAAACACAAACCA 1523
QY 2745 GCCTCACTTACACTTGTACTCAAAAGACAAAGTCAATGTCTTGTCTAGCCTT 2804
Db 1524 GCCTCACTTACACTTGTACTCAAAAGACAAAGTCAATGTCTTGTCTAGCCTT 1583
QY 2805 TTGGAAGAGGAAACAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2864
Db 1584 TTGGAAGAGGAAACAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1643
QY 2865 CACAATGAAGATATTTGTCTTTTGAATTTGTGTATTAAGCAATCAATTAATTCAG 2924
Db 1644 CACAATGAAGATATTTGTCTTTTGAATTTGTGTATTAAGCAATCAATTAATTCAG 1703
QY 2925 CAAATGATTCCTTTCTTT-AAAAAATATGTGAGGAAATTAATTTTCAAGGT 2983
Db 1704 CAAATGATTCCTTTCTTTAAAAAATATGTGAGGAAATTAATTTTCAAGGT 1763
QY 2984 TGTGGCCCAAGGCGTAAATTCAGATTTTAAAGGAAACACACAGAAAGG 3043
Db 1764 TGTGGCCCAAGGCGTAAATTCAGATTTTAAAGGAAACACACAGAAAGG 1823
QY 3044 TACCTAGGTTTAACTCAAGCACTTCTGTGTCTTGTAGATTTGTAA 3103
Db 1824 TACCTAGGTTTAACTCAAGCACTTCTGTGTCTTGTGTAGATTTGTAA 1883
QY 3104 GCTGATAGTGAAGCAATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTA 3163
Db 1884 GCTGATAGTGAAGCAATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTA 1943
QY 3164 TATCACTGCGAGCCGGAAGAGTGAAGTCAAGTGTCAACAGCTGTCTGAATTTG 3223
Db 1944 TATCACTGCGAGCCGGAAGAGTGAAGTCAAGTGTCAACAGCTGTCTGAATTTG 2003
QY 3224 CTTCCTGAGCCAAAGACAAATATGAGCTTTTGTGACAAACCTTGAATTTT 3283
Db 2004 CTTCCTGAGCCAAAGACAAATATGAGCTTTTGTGACAAACCTTGAATTTT 2063

RESULT 9
AAS70981
ID AAS70981 standard; cDNA; 2010 BP.
XX AAS70981;
AC AAS70981;
XX 13-FEB-2002 (first entry)
DT DNA encoding novel human diagnostic protein #6785.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW


```

QY 1286 ACATCAACACAGCTGGAGATTCGCCGAGCTCTATCAAGATTGCCCTGCGAAGGCC 1345
Db 1392 ACATCAACACAGCTGGAGATTCGCCGAGCTCTATCAAGATTGCCCTGCGAAGGCC 1451
QY 1346 CAGAGCTGAGGAAAGATGGTCAATCAACCGGGCCGACCGGAAAGCCAGTTCAAGGCC 1405
Db 1452 CAGAGCTGAGGAAAGATGGTCAATCAACCGGGCCGACCGGAAAGCCAGTTCAAGGCC 1511
QY 1406 AGGACGAGATCTTTGGGAACTGAAAGAGAAAATTCTTTAACTCCCAAGAGAGTGA 1465
Db 1512 AGGACGAGATCTTTGGGAACTGAAAGAGAAAATTCTTTAACTCCCAAGAGAGTGA 1571
QY 1466 AGCTGGAAGCCGATATCAAGATGCCCTCTTCCACAGCTGGCCGGGTGATTGGCAAGGTTG 1525
Db 1572 AGCTGGAAGCCGATATCAAGATGCCCTCTTCCACAGCTGGCCGGGTGATTGGCAAGGTTG 1631
QY 1526 GCAAGACGCTGAAGCACTGCAAGACTTAACCAAGTCAAGTCAATCGTCTGCTGAGAC 1585
Db 1632 GCAAGACGCTGAAGCACTGCAAGACTTAACCAAGTCAAGTCAATCGTCTGCTGAGAC 1691
QY 1586 AAACCCGAGATGAAATGAGAAATGATCGTCAAGATTATCGGCACTTTCTTGTAGCC 1645
Db 1692 AAACCCGAGATGAAATGAGAAATGATCGTCAAGATTATCGGCACTTTCTTGTAGCC 1751
QY 1646 AGACTGCAACAGCGCAAGATCAGGGAAATTTGTCACACAGTGAAGAGAGAGAGAGAAAT 1705
Db 1752 AGACTGCAACAGCGCAAGATCAGGGAAATTTGTCACACAGTGAAGAGAGAGAGAAAT 1811
QY 1706 ACCCTCAGGGAGTGGCTCAGACGCGAGCAAGTGAAGTCCCAAGGCAAGCAAGCAAAACA 1765
Db 1812 ACCCTCAGGGAGTGGCTCAGACGCGAGCAAGTGAAGTCCCAAGGCAAGCAAGCAAAACA 1871
QY 1766 ACGATGATGTAAGCCCTTCCAAACCTTGAAGATGAGACCAAGCGACCGCAAGT 1825
Db 1872 ACGATGATGTAAGCCCTTCCAAACCTTGAAGATGAGACCAAGCGACCGCAAGT 1931
QY 1826 CCGGAGCAAAACCAAGAACATCTGAGAGAAATGAGAACTGCGGAGGCGGCAAGGACTCT 1885
Db 1932 CCGGAGCAAAACCAAGAACATCTGAGAGAAATGAGAACTGCGGAGGCGGCAAGGACTCT 1991
QY 1886 GCCGAGGCCCTGAGAACCC 1904
Db 1992 GCCGAGGCCCTGAGAACCC 2010

RESULT 10
ID ACA90176 standard; cDNA; 1707 BP.
AC ACA90176;
XX
AC ACA90176;
XX
XX 10-JUL-2003 (first entry)
XX
DE cDNA encoding novel human protein NOV14a.
XX
KW Human; cytosolic; DAPK3-Agonist; DAPK3-Antagonist; cancer; NOV; gene;
XX ss.
XX Homo sapiens.
OS
XX
PN W02003031571-A2.
XX
PD 17-APR-2003.
XX
XX 02-OCT-2002; 2002WO-US031357.
XX
XX 05-OCT-2001; 2001US-0327454P.
XX 09-OCT-2001; 2001US-0327917P.
XX 09-OCT-2001; 2001US-0328029P.
XX 09-OCT-2001; 2001US-0328056P.
XX 12-OCT-2001; 2001US-032849P.
XX 15-OCT-2001; 2001US-0329414P.

```

```

PR 17-OCT-2001; 2001US-0330142P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0349357P.
PR 25-JUN-2002; 2002US-0391342P.
PR 01-OCT-2002; 2002US-00262445.
XX
XX (CUBA-) CUBAGEN CORP.
XX
XX Alsebrook JP, Burgess CE, Catterton E, Chant JS, Chaudhuri A;
XX Edinger SR, Gerlach VL, Giot L, Gorman L, Guo X, Kekuda R;
XX Mezes PS, Millet I, Ooi CE, Patnubajan M, Rieger DK, Szytek KA,
XX Taupier RJ, Zernusen BD, Zhong H, Zhong M;
XX
XX MPI; 2003-381704/36.
XX
XX P-PsDB; AB089799.
XX
XX New DAPK3 polypeptide, useful for preparing a composition for treating or
XX preventing e.g., cancer.
XX
XX Claim 20; Page 129; 253pp; English.
XX
XX The invention describes an isolated polypeptide comprising any of 33 90-
XX 1273 amino acid sequences (I) given in the specification or its mature
XX form, a sequence that is at least 95 % identical to (I), or a sequence
XX comprising one or more conservative substitutions in the amino acid
XX sequence of (I). The polypeptide is useful for preparing a composition
XX for treating or preventing e.g. cancer. This sequence encodes a novel
XX human NOV protein
XX
XX Sequence 1707 BP; 492 A; 443 C; 442 G; 330 T; 0 U; 0 Other;
SQ
Query Match 47.6%; Score 1561.2; DB 7; Length 1707;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 83; Indels 3; Gaps 1;
QY 66 ACGATGATGAACAAGCTTTAATGAGGAACTGAGCCCGCGCTACCGCGAGAGCCT 125
Db 1 ACGATGATGAACAAGCTTTTATCGGGAACCTGAGCCCGCGCTACCGCGAGAGCCT 60
QY 126 CCGGAGCTCTTTGGGAGAGAGAGTGCCTCGGCGGAGCAAGTCTGTAAGTCCG 185
Db 61 CCGGAGCTCTTTGGGAGAGAGAGTGCCTCGGCGGAGCAAGTCTGTAAGTCCG 120
QY 186 CTAAGCCTCGTGAATAACCCGACCAAGAACTGGGCATCCGGCCATGAGACCTCTC 245
Db 121 CTAAGCCTCGTGAATAACCCGACCAAGAACTGGGCATCCGGCCATGAGACCTCTC 180
QY 246 GGGTAAAGTGAATTCATGAGGAAATCATGGAAGTTGATTACTGAGTCTTAAAGCT 305
Db 181 GGGTAAAGTGAATTCATGAGGAAATCATGGAAGTTGATTACTGAGTCTTAAAGCT 240
QY 306 AAGGACAGGAAATTCAGATTGAAACATCCCTCTCACTCACTGAGGAGGTGTTGA 365
Db 241 AAGGACAGGAAATTCAGATTGAAACATCCCTCTCACTCACTGAGGAGGTGTTGA 300
QY 366 TGAATTTTGGCTCAATATGAGACAGTGAAGATGTTGAACAATGCAACAGACACAGA 425
Db 301 TGAATTTTGGCTCAATATGAGACAGTGAAGATGTTGAACAATGCAACAGACACAGA 360
QY 426 AACCGCGTTGTCAGCTCAATATGCAACAAGAGAAAGCAAAAATAGCATGAGAA 485
Db 361 GACCGCTGTTGTCAGCTCAATATGCAACAAGAGAAAGTAAATAATAGCATGAGAA 420
QY 486 GCTAAGGGGCAATCAGTTGAGAACTACTCTTCAAGATTCTTCACTCCGAGTGAAGA 545
Db 421 GCTAAGGGGCAATCAGTTGAGAACTACTCTTCAAGATTCTTCACTCCGAGTGAAGA 480
QY 546 GGTAGCTCCCTTGGCCCTTCAAGCAGAGCCAGCGTGGGAGCACTTTCCGGAGAGA 605
Db 481 GGTAGCTCCCTTGGCCCTTCAAGCAGAGCCAGCGTGGGAGCACTTTCCGGAGAGA 540

```


QY	1307	TCGCCGAGGCTCTATCAAGATTGCCCTTCGCGAAGGCCAGACGTCCGCGAAGAGATG	1366
Db	85	TCGCGGAGGCTCTATCAAGATTGCCCTTCGCGAAGGCCAGACGTCCGCGAAGAGATG	144
QY	1367	TCATCATCACCGGGCCACCGGAAAGCCCACTTCAAGGCCAGGGAAGGATCTTTGGGAAC	1426
Db	145	TCATCATCACCGGAGCCA - CGSAAAGCCCACTTCAAGGCCAGGGAAGGATCTTTGGGAAC	203
QY	1427	TGAAAGAGGAAACTCTTTAAACCCCAAGAGAGATGAAGCTGGAACGCAATATCAAG	1486
Db	204	TGAAAGAGGAAACTCTTTAAACCCCAAGAGAGATGAAGCTGGAACCAATATCAAG	263
QY	1487	TGCGCTCTTCCACAGCTGCGCGGATGTGGCAAAAGTGGCAAGACCGTGAACCACTGC	1546
Db	264	TGCGCTCTTCCACAGCTGCGCGGATGTGGCAAAAGTGGCAAGACCGTGAACCACTGC	323
QY	1547	AGAACTTAAACAGTGCAGAAAGTCATCTGTGCTCTGTACCAACCGCAGATGAAATGAG	1606
Db	324	AGAACTTAAACAGTGCAGAAAGTCATCTGTGCTCTGTACCAACCGCAGATGAAATGAG	383
QY	1607	AAGTATCGTCAGAAATTAATCGGGCACTCTTTGCTAGGCAACATGCAACAGGCAAGATCA	1666
Db	384	AAGTATCGTCAGAAATTAATCGGGCACTCTTTGCTAGGCAACATGCAACAGGCAAGATCA	443
QY	1667	GCGAAATTTG - TACAA CAGGTGAAGCAGAGAGCAAGAAATACCTTCAGGAGTGCCTCA	1725
Db	444	GCGAAATTTGTTACAA CAGGTGAAGCAGAGAGCAAGAAATACCTTCAGGAGTGCCTCA	503
QY	1726	CAGCCCAAGAAATGAGGCTCCCA CAGGCA CCAAGCAAA CAA CAGATGAATGATCCCTTC	1785
Db	504	CAGCCCA - NAAATGAGGCTCCCA CAGGCA CCAAGCAAA CAA CAGATGAATGATCCCTTC	562
QY	1786	CAACACCTGACAGATGAGCAACCAACGACGACGACGATGGGAGGCAACCAAGACCA	1845
Db	563	CAACACCTGACAGATGAGCAACCAACGACGACGACGATGGGAGGCAACCAAGACCA	622
QY	1846	TCTGAGAGTAGAAGTGTGCGAGGCGGCGAGGAGACTGTGCGAGGCTCTGAGAACCC	1905
Db	623	TCTGAGAGTAGAAGTGTGCGAGGCGGCGAGGAGACTGTGCGAGGCTCTGAGAACCC	682
QY	1906	AGGGGCGAGAGAGGGGCGGGGAAGGTCA GCGAGGTTTGCAGAACCA CCGAGCCCGCT	1965
Db	683	AGGGGCGAGAGAGGGGCGGGGAAGGTCA GCGAGGTTTGCAGAACCA CCGAGCCCGCT	742
QY	1966	CCCGGCCCCCAGGGGCTTTCGAGGGCTTCAGGCAAT - CCACTTCAACATCTCACTGGATCTC	2024
Db	743	CCCGGCCCCCAGGGGCTTTCGAGGGCTTTCAGGCAATCTTCAGATCACTCGATCTC	802
QY	2025	TCTGAACTCCCAAGACGCTATCCCTTTAGTTAGTTAACTAACTATAGTGAACGTGTCAAA	2084
Db	803	TCTGAACTCCCAAGACGCTATCCCTTTAGTTAGTTAACTAACTATAGTGAACGTGTCAAA	862
QY	2085	GCCAAAGCAAAATGCA CACCTTTTCTGTGCAAAATGCTCTGTACATGTGTATACATA	2144
Db	863	GCCAAAGCAAAATGCA CACCTTTTCTGTGCAAAATGCTCTGTACATGTGTATACATA	922
QY	2145	TTAGAAAGGGAAGAGTGAATATGTGTGCGCTGTGGGTACACAAGGATGCTCAGCGGT	2204
Db	923	TTAGAAAGGGAAGAGTGAATATGTGTGCGCTGTGGGTACACAAGGATGCTCAGCGGT	982
QY	2205	AATATATTTTGAATATATATATCAATTAATCACTCACTCACTCACTCACTCACTCACT	2264
Db	983	AATATATTTTGAATATATATATCAATTAATCACTCACTCACTCACTCACTCACTCACT	1042
QY	2265	TAAATTTTTTTTCTTTTAAAGAGAAAGCGGCTTTCTAGACTTTAAAGATTAAGTCT	2324
Db	1043	TAAATTTTTTTTCTTTTAAAGAGRAAGCGGCTTTCTAGACTTTAAAGATTAAGTCT	1102
QY	2325	TTGGAGAGTCA CAGGTGTGAGAGAGGCTTTGAGGCA CCGGCA CAAATTTCA CCGCA	2384
Db	1103	TTGGAGAGTCA CAGGTGTGAGAGAGGCTTTGAGGCA CCGGCA CAAATTTCA CCGCA	1162
QY	2385	GGAATATCTGTCGAGAGACACT 2408	

Db	1163	GGGAATCTCTGTCGGAAGACACT	1186
RESULT 12			
ABX73911			
ID	ABX73911	standard; DNA; 1186 BP.	
XX			
AC	ABX73911;		
XX			
DT	18-MAR-2003	(first entry)	
XX			
DE	Human novel polynucleotide #739.		
XX			
KM	Human; gene; ds; neural disorder; immune system disorder; renal disorder		
KM	muscular disorder; respiratory disease; reproductive disorder;		
KM	gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;		
KM	hyperproliferative disorder; inflammatory disease; allergic reaction;		
KM	blood related disorder; cancer; immunosuppressive; antiinflammatory;		
KM	cardiovascular; nephrotoxic; cytostatic; antiallergic; thrombolytic;		
KM	haemostatic; antiarteriosclerotic.		
XX			
OS	Homo sapiens.		
XX			
PN	US2002132753-A1.		
XX			
PD	19-SEP-2002.		
PF	17-JAN-2001; 2001US-00764864.		
XX			
XX	31-JAN-2000; 2000US-0179065F.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225457P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	22-AUG-2000; 2000US-0225758P.		
PR	30-AUG-2000; 2000US-0226868P.		
PR	01-SEP-2000; 2000US-0228924P.		
PR	01-SEP-2000; 2000US-0229287P.		
PR	01-SEP-2000; 2000US-0229343P.		
PR	01-SEP-2000; 2000US-0229344P.		
PR	01-SEP-2000; 2000US-0229345P.		
PR	05-SEP-2000; 2000US-0228509P.		
PR	05-SEP-2000; 2000US-0229513P.		
PR	08-SEP-2000; 2000US-0231413P.		
PR	21-SEP-2000; 2000US-0234223P.		
PR	21-SEP-2000; 2000US-0234274P.		
PR	25-SEP-2000; 2000US-0234997P.		
PR	27-SEP-2000; 2000US-0235834P.		
PR	29-SEP-2000; 2000US-0236327P.		
PR	29-SEP-2000; 2000US-0236367P.		
PR	29-SEP-2000; 2000US-0236368P.		
PR	29-SEP-2000; 2000US-0236369P.		
PR	29-SEP-2000; 2000US-0236370P.		
PR	02-OCT-2000; 2000US-0236802P.		
PR	02-OCT-2000; 2000US-0237037P.		
PR	02-OCT-2000; 2000US-0237038P.		
PR	02-OCT-2000; 2000US-0237039P.		
PR	02-OCT-2000; 2000US-0237040P.		
PR	13-OCT-2000; 2000US-0239935P.		
PR	20-OCT-2000; 2000US-0240960P.		

PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251866P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-14744/14.
DR P-PSDB; ABUS5651.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 749; 402bp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 1186 BP; 338 A; 310 C; 302 G; 230 T; 0 U; 6 Other;
Query Match 33.4%; Score 1096.8; DB 7; Length 1186;
Best Local Similarity 98.4%; Pred. No. 6.8e-281;
Matches 1145; Conservative 3; Mismatches 12; Indels 4; Gaps 4;
QY 1247 CCCAGGCTGTGGCGCCATCATCGGAGAGAGGGGCGACACATCAACAGCTGGCGAGAT 1306
DB 25 CCCAGCGTCCGCGCCATCATCGGAGAGAGGGGCGACACATCAACAGCTGGCGAGAT 84
QY 1307 TGGCGGAGCTTATCAAGATTGCCCTGCGGAGAGGCCCAAGCGTCAAGAGAGATG 1366
DB 85 TGGCGGAGCTTATCAAGATTGCCCTGCGGAGAGGCCCAAGCGTCAAGAGAGATG 144
QY 1367 TCATCAATCCGCGGAGAGGCCCAAGTTCAGGAGGCCGAGATCTTTGGAAAC 1426
DB 145 TCATCAATCCGCGGAGAGGCCCAAGTTCAGGAGGCCGAGATCTTTGGAAAC 203
QY 1427 TGAAGAAGGAACTCTTTACCCCAAGAGAGAGTGAAGTGAAGCCCATATCAAG 1486
DB 204 TGAAGAAGGAACTCTTTACCCCAAGAGAGAGTGAAGTGAAGCCCATATCAAG 263
QY 1487 TGGCCCTTTTCAAGCTGGCGGCGGTATTTGGCAAAAGTGGCAAGCCGTGAAGAACTGC 1546
DB 264 TGGCCCTTTTCAAGCTGGCGGCGGTATTTGGCAAAAGTGGCAAGCCGTGAAGAACTGC 323
QY 1547 AGAAGTTAACGAGTGAAGTCAATGCTGCTGCTGCAACAAAGCCAGATGAATAGG 1606
DB 324 AGAAGTTAACGAGTGAAGTCAATGCTGCTGCTGCAACAAAGCCAGATGAATAGG 383
QY 1607 AAGTGATCGTCAAGATTAATCGGCACTTTCTTTAGTACAGATGCAAGGCGCAAGTCA 1666
DB 384 AAGTGATCGTCAAGATTAATCGGCACTTTCTTTAGTACAGATGCAAGGCGCAAGTCA 443

QY 1667 GGGAAATTG-TACAAAGGTGAGAGAGAGAGAGAGAAATACCTTCAAGAGTGCCTCA 1725
DB 444 GGGAAATTGTTACAAAGGTGAGAGAGAGAGAGAGAAATACCTTCAAGAGTGCCTCA 503
QY 1726 CAGCGAGCAATGAGGCTCCCAAGAGCAACAGGAAACAAAGATGAATGAGCCCTTC 1785
DB 504 CAGCGCA-NAAGTGAAGCTCCCAAGAGCAACAGGAAACAAAGATGAATGAGCCCTTC 562
QY 1786 CAACACCTGACAGAAATGAGAGACCAAGCAGCAGCAGATCGGAGAGCAACCAAGCA 1845
DB 563 CAACACCTGACAGAAATGAGAGACCAAGCAGCAGCAGATCGGAGAGCAACCAAGCA 622
QY 1846 TCTGAGAAATGAGAGATCTGCGAGAGCGCGGAGAGAGTCTTGCAGAGCCCTGAGAACCC 1905
DB 623 TCTGAGAAATGAGAGATCTGCGAGAGCGCGGAGAGAGTCTTGCAGAGCCCTGAGAACCC 682
QY 1906 AGGGGCGGAGAGAGGCGGAGAGAGTCAAGGTTTSCCAAGCCAGAGCCCGGCT 1965
DB 683 AGGGGCGGAGAGAGGCGGAGAGAGTCAAGGTTTSCCAAGCCAGAGCCCGGCT 742
QY 1966 CCGGCGCCCGCAGGCGCTTCTGAGAGCTTCAAGCAT-CCACTTCAACCATCGATGATCTC 2024
DB 743 CCGGCGCCCGCAGGCGCTTCTGAGAGCTTCAAGCAT-CCACTTCAACCATCGATGATCTC 802
QY 2025 TCTGAACTCCCAAGAGCTATCCCTTTAGTTAGTAACTATAGTGAAGCTTTCAA 2084
DB 803 TCTGAACTCCCAAGAGCTATCCCTTTAGTTAGTAACTATAGTGAAGCTTTCAA 862
QY 2085 GCGAAGCAAAATGCAACCCCTTTCTGAGCAAAATGCTCTGTAATGATGTGATATA 2144
DB 863 GCGAAGCAAAATGCAACCCCTTTCTGAGCAAAATGCTCTGTAATGATGTGATATA 922
QY 2145 TTAGAAAGGAAAGTATAGATATGAGCTGTGAGTATCAAGAGTCTGAGAGGCT 2204
DB 923 TTAGAAAGGAAAGTATAGATATGAGCTGTGAGTATCAAGAGTCTGAGAGGCT 982
QY 2205 AATATATTTTAAAGTAAATATATATCAATATCAATCAATCAATCAATCAAT 2264
DB 983 AATATATTTTAAAGTAAATATATATCAATATCAATCAATCAATCAATCAAT 1042
QY 2265 TAATTTTTTTCTTTTAAAGAGAGAGAGGCTTTTCAAGCTTAAAGATTAAGTCT 2324
DB 1043 TAATTTTTTTCTTTTAAAGAGAGAGAGGCTTTTCAAGCTTAAAGATTAAGTCT 1102
QY 2325 TTGGAGAGCTCAAGGCTTGAAGAGAGAGGCTTTGAAGCCACCGCAAAATTCACCA 2384
DB 1103 TTGGAGAGCTCAAGGCTTGAAGAGAGAGGCTTTGAAGCCACCGCAAAATTCACCA 1162
QY 2385 GGGAAATCTGTCGGAAGACACT 2408
DB 1163 GGGAAATCTGTCGGAAGACACT 1186
RESULT 13
AAC66035
ID AAC66035 standard; cDNA; 1740 BP.
XX
AC AAC66035;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human lung cancer-associated cDNA antigen L523S.
XX
KW lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW vaccine; detection; ss.
XX
OS Homo sapiens.
XX
PN WO200061612-A2.
XX
PD 19-OCT-2000.
XX

PF 03-APR-2000; 2000MO-US008896.
XX
PR 02-APR-1999; 99US-00285479.
PR 17-DEC-1999; 99US-00466396.
PR 30-DEC-1999; 99US-00476496.
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Fan L,
XX
DR WPI, 2000-628399/60.
DR P-PSDB; AAB11365.
XX
PT Isolated polypeptide comprising an immunogenic portion of a lung tumor
PT protein is used for detecting and monitoring progression of lung cancer
PT in a patient.
XX
PS Claim 1a; Page 258-259; 261p; English.
XX
CC This invention describes a novel isolated polypeptide (I) which
CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
CC which have cytostatic activity. The polypeptides and polynucleotides are
CC used in compositions and vaccines to inhibit the development of cancer,
CC especially lung cancer, in a patient. Methods described in the invention
CC can be used to monitor the progression of a cancer by carrying out the
CC detection at subsequent time points and comparing the results from the
CC different time points. CD4+ and/or CD8+ T-cells isolated from a patient
CC are treated with P2, polynucleotides encoding P2 or antigen presenting
CC cells expressing P2 and then administered to the patient to inhibit
CC development of cancer
XX
SQ Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;

Query Match 21.2%; Score 697.2; DB 3; Length 1740;
Best Local Similarity 64.9%; Pred. No. 2.1e-174;
Matches 1116; Conservative 0; Mismatches 538; Indels 66; Gaps 3;

QY 73 ATGAAACAAGCTTTATCATCGGGAACCTGAGCCCGCCGTCACCGCCGACGACTTCGCGCAG 132
DB 1 ATGAAACAAGCTTTATCATCGGGAACCTGAGCCCGCCGTCACCGCCGACGACTTCGCGCAG 60

QY 133 CTCTTTGGGGACAGAAAGCTGCCCCCTGGCGGGACAGGTCCTGTGAAGTCCGGGTAGGCC 192
DB 61 ATCTTCAAGACGCGCAAGATCCCGGTGTGCGGACCCCTTCTGTGAAGACTGGCTACGCG 120

QY 193 TTGCTGACCTACCCCGACAGAACTGCGCCATTCGCGCCATCGAAGCCCTCTCGGTTAA 252
DB 121 TTCTGTGACCTGCCCGGACGAGAGCTGGGCTCTCAAGCCATCGAGCGCTTTCAGTTAA 180

QY 253 GTGGAATTCATGGGAATATCATGGAGTTGATTAATCATGCTTTAAAGCTTAAGAGC 312
DB 181 ATGAAATCTGACGCGAAACCCATGAAAGTTGAGCACTGCGTCCCAAAAGCAAGATT 240

QY 313 AGGAAATTCAGATTGGAACATCCCTCTCACTGAGTGGGAGGTGTTGATGACTT 372
DB 241 CGGAAATTCAGATTGGAATATCCCTCTCTTAAAGTGGAGGTGCTGATAGTTTA 300

QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGAAACAACTCAACAGACAGAAACCGCC 432
DB 301 CTAGTCCAGATAGAGAGTGTGAGAGCTGTGAGCAATGAACTGACTGCGAAACAGCA 360

QY 433 GTTGTCAAGCTCATATTTGCAACAAGAGAAAGCAAAATAGCCATGGAAGCTAAGC 492
DB 361 GTTGTAAATGTAACTTATTCAGTAAAGCAAGCTAAGACAACTGAATCTGAAT 420

QY 493 GGGCATCAGTTTGAAGACTACTCTCAAGATTCTCTACATCCCGATGAAGAGGTGAGC 552
DB 421 GGAATTCAGTTAAGAAATTCACCTTGAAGAGTACCTTATCTCTGATGAAGAGCGCGCC 480

QY 553 TCCCTTCGCGCCCTCAGGAGCGCCAGCTGGGGAACAATCTTCTCCGGAGCAAGGC--- 609

DB 481 CAGCAAAACCCCTTGACAGAGCCCCGAGAGTGCCTGGGGCTTGGGACAGAGGCTCTCA 540
QY 610 -----CAGCGCCCTGGGGGACATTCTCAGGCGACAGATGATTTCCCGTGGGATC 663
DB 541 AGGCAGAGGCTCTCCAGGATCCGATTCGAAGAGAAACATGATTTGCTGTGGCGCTG 600
QY 664 CTGATCCCAACCAATTTGTTGTGTGCTCATCGGAAAGAGGCTTGACCATTAAGAAC 723
DB 601 CTGTTCCCAACCAATTTGTTGTGAGCCATCATAGGAAAGAGGTGCCCATTTGGAAAC 660
QY 724 ATCACTAAGCAGACCCAGCTCCCGGTAGATATCATAGAAAGAGAACTGTGAGGTGA 783
DB 661 ATCAACCAACAGACCCAGCTTAATATCATATGTCACCTTAAGAAATGCGGGGCTGT 720
QY 784 GAGAAAGCTGTACCATTCATGCCACCCGAGAGGAGCTTGAAGATGCGCATATT 843
DB 721 GAGAAAGTGAATTAATCTCTCTCTATCTCTGAAGCACTTCGCGCTTGAAGTCTATT 780
QY 844 CTGAAATCATGACAGAAAGAGGCAATGAGACCAACTAGCCGAGATTCCTTGAAA 903
DB 781 CTGAGATTTATGCAATMAAGAACTCAGATTAATAATTCACAGAAAGATCCCTTGAA 840
QY 904 ATCTTGGCACCAATGCTGTGTGGAAGATGGAAGTGAAGAAAGAGCAAAATTTGAAG 963
DB 841 ATTTTAGCTAATTAATCTTTGTTGAGCTCTTATTGTAAAGAAAGAAATCTTAA 900
QY 964 AAAATTGAAACATGAAACAGAGGACCAAGATTAACAATCTCATCTTTGACAGATTGACATA 1023
DB 901 AAAATTGAGCAAGACAGACATTAATAATCAAGATATCCATTTGACAGAAATTGACGCTG 960
QY 1024 TACAAACCGGAAAGAACATTAATCTGTAAGGACAGCTTGAAGGCTGTGCAAGTCTGAG 1083
DB 961 TATATCCGAAACCCATTAATTAATCAAGTAAAGCAATGTTGAGATGTCGCAAGAGCTGAG 1020
QY 1084 ATAGAGATTTAAGAAAGCTGCTGAGGCTTTGAAATGATATGCTGCTGTTAACACC 1143
DB 1021 GAGAGATTCATGAAGAAATATGAGGAGCTTATGAATGATTTGCTTATGATATCTT 1080
QY 1144 CACTCGGAT-----ACTTCCCA 1162
DB 1081 CAAGCATTAATTAATCTCGATTAATCTGAAGCCTTGGTCTGTTCCCAACCACTTCA 1140
QY 1163 GCGGTACCCCATCACCAGTTTGGCCGTTCCCGCATCATCTC----- 1208
DB 1141 GGAATGCCACTCCCACTCAGGCGCCCTTCACACCACTGACTCCCTCAACCCGAGTTT 1200
QY 1209 --TTATCAGAGCAGAGATTTGATCTCTCATATCCCAACCCAGGCTGTGGGCGCATC 1266
DB 1201 GAGAAATCAGAAACGAGAGCTGTTCACTGTTTATCCAGCTTATCAAGTGGTGCATC 1260
QY 1267 ATGGGAAAGAGGAGGACATCAACCACTGCGAGATTCGCGGAGGCTCTATCAAG 1326
DB 1261 ATCGGCAAGCAGGCGCAGCATCAAGCACTTCTGCTTGTGAGAGCTTCAATTAAG 1320
QY 1327 ATTGCCCTCGGAAAGGCCCAAGCTCAGAGTGAAGAAAGATGTGTCATCATCCGCGCACCG 1386
DB 1321 ATTGCTCAAGCGAAGACCAAGATGCTTAAAGTGAAGTGTGATTAATCACTGACACCA 1380
QY 1387 GAAGCCAGTTCAAGGCCGAGAGCGAGATTTTGGGAACTGAAAGAGAAACTCTTT 1446
DB 1381 GAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGAAAGAAATTAAGAGAAACTTTGTT 1440
QY 1447 AACCCCAAGAAAGATGAAAGCTGGAAGGCAATATGAGATGCTCTTCAAGAGTGC 1506
DB 1441 AGTCTTAAAGAAAGAGTGAAGTGAAGCTCAATCAGATGTCATCTTGTGCTGCGC 1500
QY 1507 CGGGTATTTGGCAAGGTGCAAGACCGTGAAGAACTGAGAACTTAACAGTGCAGAA 1566
DB 1501 AGAGTTATGGAAGAAAGAGGCAAAAGGTGAAGAACTTCAAGATTTGTCAGAGTGA 1560
QY 1567 GTCATGCTGCTGTGACCAAGCGCAGATGAAATGAGGAAAGTATCGCAATTAATC 1626
DB 1561 GTTGTGTCTCTGTGACAGACACTGATGAGATGACCAAGTGTGTCATAAATTAATC 1620


```
Db 1141 GGGATGCCA|CTCCCACTCAGGCGCC|CCCTCAGCATATCTCTCCCTTACC|CGAGTTT 1200
Qy 1209 --TTATCCAGAGAGATGTGATCTCTTCATCCCA|CCAGGCTGTGGCCG|CATC 1266
Db 1201 GAGCATCA|GAAACGAGACTGTTCATCTGTTATCCCA|GCTCTATCATCTGCTG|CCATC 1260
Qy 1267 ATCGGGAAGAA|GGGGGCAACATCAAA|CAGCTGGGAGATTCGGCGGAG|CCCTTATCAAG 1326
Db 1261 ATCGGCAAGAG|GGGCGAGACATCAAGACATCTTCGCTTGTCTGGAGCTTCA|TTAAG 1320
Qy 1327 ATTGCCCC|TGGGGAAGGCCGACGTCACGGAAGATGTGTCATCATCA|CCGGGCGACCG 1386
Db 1321 ATTGCTCAG|CGGAAAGCACACAGATCTTAAGTGAAGATGTGATATATCA|CTGAGCAACA 1380
Qy 1387 GAAGCCCA|GTTCAAGSCCAGGAGGATCTTTGGGAAATCGAAAGAGAA|ACTTCTTT 1446
Db 1381 GAGGCTCA|GTTCAAGGCTCAGGGAAGATTTATGGAAAAATTTAAAGAA|AAACTTTGTT 1440
Qy 1447 AACCCAA|AAGAAAGTGAAGCTGGAAGCGCATATCAGATG|CCCTTCCACAGCTGGC 1506
Db 1441 AGTCCTAA|AAGAAAGGTGAATCTTGAAGCTCATATCAGATG|CCATCTTGTGCTGGC 1500
Qy 1507 CGGGTGA|TTGGCAAAAGTGGCAAGACGCTGAACGA|ACTGCAAACTTAACGATG|CAGAA 1566
Db 1501 AGAGTTA|TTGGAAAAAGAGGCAAAA|CGGTGAATGAATCTTCA|GAATTTGTCAAGT|GCAAA 1560
Qy 1567 GTCAATCG|CTCGTCCGACCAACCGCATGAAATGAAATG|GAAAGTATGTGATCA|AATATC 1626
Db 1561 GTTGTGT|CTCCTCGTACGACACACTGATGAGATGA|CCAAAGTGTGTGCA|AAATTA|ACT 1620
Qy 1627 GGGCACTT|CTTTTGAAGCAGACAGCAGCGCAAGAT|CAGGAAATTTGTACAA|CAGGTTG 1686
Db 1621 GGTACTT|CTATGATGTGCGAGGTGGCCAGAGAA|AAATTCAGGAAATTTCTGA|CTCAGGTA 1680
Qy 1687 AAGCAGA|GAGACGAAATATACCTCAGGAGTCCCTC|AC 1726
Db 1681 AAGCAGA|CCAAACAGAAAGCTCTGCAAA|GTGAGCACAC 1720

RESULT 15
ID ABQ92440 standard; cDNA; 1740 BP.
XX
AC ABQ92440;
XX
DE 07-OCT-2002 (first entry)
XX
DE Human lung cancer associated cDNA sequence SEQ ID NO:347.
XX
KW Human; lung cancer; lung tumor; cytostatic; gene therapy; vaccine; gene;
XX ss.
XX
OS Homo sapiens.
XX
PN WO200247534-A2.
XX
PD 20-JUN-2002.
XX
PF 30-NOV-2001; 2001MO-US047576.
XX
PR 12-DEC-2000; 2000US-00735705.
XX
PR 07-MAY-2001; 2001US-00850716.
XX
PR 28-JUN-2001; 2001US-00897778.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI Monelli PD, Fanger N, Retter MW, Durham M, Fanger GR, Veddyck TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR WPI; 2002-583465/62.
DR P-FSDB; ABP61917.
```

```
XX
PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT the polynucleotides, useful in pharmaceutical compositions such as
PT vaccines and as markers to indicate the presence of lung cancer.
PS Claim 1, Page 337; 381pp; English.
XX
CC The present invention describes isolated human lung carcinoma
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC activity, and can be used in gene therapy and in vaccines. Compositions
CC comprising (I) or (II) can be used for stimulating an immune response in
CC a patient and for treating lung cancer in a patient. Oligonucleotides of
CC (I) can be used for detecting the presence of a cancer in a patient, by
CC obtaining a biological sample from the patient, contacting the biological
CC sample with the oligonucleotide, detecting in the sample, an amount of
CC polynucleotide that hybridizes to the oligonucleotide and comparing the
CC amount of polynucleotide that hybridizes to the oligonucleotide and the
CC predetermined cut-off value, and determining the presence of a cancer in
CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC vaccines. (I) is useful as a marker to indicate the presence or absence
CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC ABP61992 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 1740 BP; 526 A; 406 C; 417 G; 391 T; 0 U; 0 Other;
Query Match 21.2%; Score 697.2; DB 6; Length 1740;
Best Local Similarity 64.9%; Pred. No. 2.1e-174;
Matches 1116; Conservative 0; Mismatches 538; Indels 66; Gaps 3;
Qy 73 ATGAA|CAAGCTTTA|CATTCGGGA|ACCTGAG|CCCGCCGTCAC|CGCGAC|CTCCGCGAG 132
Db 1 ATGAA|CAAGCTTTA|CATTCGGGA|ACCTGAG|CAAGCA|AGCCCGCCCTCGAC|CTTAAAGT 60
Qy 133 CTCTTTGGG|GACAGAAAGTGC|CCCTGGCGGGA|AGATTCCTG|GAGTCCGGTAC|CGCC 192
Db 61 ATCTTCA|AGACGCGCAAGATCC|GGTTCGGGAC|CTTCTCGTGA|AGACTG|GGCTAC|CGC 120
Qy 193 TTGCTGAG|TACCCCGACAGAA|CTGGGCCAT|CCGCCAT|CGAC|CCCTTCGGGTAA 252
Db 121 TTGCTGAG|TACCCCGACAGAGAGTGG|GCCCTCAAG|GCCATCGAG|GGCTTTCAG|GTAA 180
Qy 253 GTGGA|ATTGATGGGAAATATCA|TGAAGTGTATCTCA|GTTCTTA|AAAAGCTA|AGAGC 312
Db 181 ATGAA|ACTGACGCGAAACC|CATAGAA|GTGAGCA|CTCGGTCC|CAAAAAGCA|AGGATT 240
Qy 313 AGGAA|AATTCAGATTG|GAAACATCC|CTC|ACCTGAG|GGGAGG|GTGGA|TGACTT 372
Db 241 CGGAA|AATTCAGATTG|GAAATATCC|CTCATTTA|CAGTGG|AGGTG|GTGATG|TTA 300
Qy 373 TTGCTCA|TATGAGGAC|GTGGAATGTG|GAAACA|AGTCA|ACAGAC|AGAAAC|CGCC 432
Db 301 CTAGTCC|AGTATGAGTGGTGG|AGAGCTG|GACAGATGA|ACTGACTG|GAAACTGCA 360
Qy 433 GTTGTCA|AGTCAATATG|CAACAAGAA|GCAAAATATG|CCATG|GAGAA|GCTA|AGC 492
Db 361 GTTGTAA|ATGTAAC|TAATTCAGTAA|GACCAAG|CTAGAC|AACTAGCA|AAACTGAAT 420
Qy 493 GGGCATCA|GTTTGA|GAACTA|CTCCTCA|GATTTC|TCACT|CCCGATG|GAAGGTAGC 552
Db 421 GGAATTC|AGTTAGAA|ATTTCA|CTTGA|AAAGT|AGCCAT|ATGCTGATG|AAAGCGCGCC 480
Qy 553 TCCCTTC|GGCCCTTC|GAGCGAC|CTTCTCAG|GCCAG|CAGATTATTT|CCGCTGCGGATC 663
Db 481 CAGCA|AAACCCCTTGGAG|CAGCCCGAGG|CGCGGGGCTTGGG|CAGAGGGGCTCTCA 540
Qy 610 -----CAGGCCCTTGGGGG|CACTTCTCAG|GCCAG|CAGATTATTT|CCGCTGCGGATC 663
Db 541 AGGCA|GCGGCTTCTCAG|ATCCGATCA|AGCA|AAACCA|TGTGATTTG|CTCGCGCTG 600
Qy 664 CTGATCC|CAACCGAGTGTG|GTGCTCAT|CGAAAGAGG|GCTTGA|CATAAAGAC 723
Db 601 CTGATCC|CAACCGAATTTGTG|AGCCATCA|TAGAA|AAAGAGTGC|CACTTTCGGAAC 660
```

QY 724 ATCCTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTCTGGAGCTGCA 783
DB 661 ATCACCAAAACAGACCCAGCTTAAATCGATGTCACCTTAAAGAAAAGCGGGGCTGCT 720
QY 784 GAGAAAGCTGTCACATCCATGATCCGCCAGAGGGGAGCTTCTGAAGCATGCCGATGAT 843
DB 721 GAGAAAGTGAATTAATCTCTCTTACTCTGAAAGGCACCTCTGGGGCTTGTAAGTAT 780
QY 844 CTGGAATCATGCAAGAAAGAGCAGATGAGACCAACTAGCCGAAGAGATTCCTGAAA 903
DB 781 CTGGAGATTAAGCAATAGAAAGCTCAAGATTAATTAATTCACAGAGAGATCCCTTGAAG 840
QY 904 ATCTTGGCAGCATAGCTGTTGTTGTTGAAGACTGATTGGAAGAAAGAGCAAAATTGAAG 963
DB 841 ATTTAGCTCAATTAATTAATCTTGTGAGCGTCTTATGTTGAAGAAAGAAATCTTAA 900
QY 964 AAAATTGAACATGAACAGAGGCAAGATTAACAATCTCATCTTTCAGAGATTGACATA 1023
DB 901 AAAATTGAGCAAGACACAGACCTAAATCAAGATATCTCCATTCAGAGAAATTGACGCTG 960
QY 1024 TACAAACCCGAAAGAACCATCACTGTGAAGGGCAAGTTAGGGCTGTGCCAGTGTGAG 1083
DB 961 TATAATCCAGAACGCACTATTAAGTTAAAGCAATGTTAGACATGTGCCAAAGCTGAG 1020
QY 1084 ATAGGATTAAGAAAGCTGCTGAGGCTTGAAGAAATGATATGCTGCTGTTAACACC 1143
DB 1021 GAGGAGATCATAGAAATAGGAGTCTTAAGAAATGATATGCTTCTATGAAATCTT 1080
QY 1144 CACTCCGGAT-----ACTTCTCCA 1162
DB 1081 CAAGACATTTAATCTTGATTAATCTGAAGGCTTGCTGTGTTCCCACTTCA 1140
QY 1163 GCCTGATCCCATCAACAGTTTGCGCTTCCCGCATCATC----- 1208
DB 1141 GGGATGCCACTCCCACTCAGGGCCCTTCAGCCATGACTCTCCCTAACCCGCAATT 1200
QY 1209 --TTATCCAGACAGAGATTTGMAATCTTCATCCCAACCAAGCTGTTGGGGCGCATC 1266
DB 1201 GAGCAATCAAGAAAGGAGCTGTTCATCTGTTATCCAGCTCTATCAGTGGTGCATC 1260
QY 1267 ATCGGGAAGAGGGGACACATCAAAACAGCTGCGAGATTGCGCCGAGCTCTATCAAG 1326
DB 1261 ATCGGAAGACAGGCGCAGACATCAAGCGCTTCTCGCTTGTGAGCTTCAATTAAG 1320
QY 1327 ATTGCCCTGCGAGAGGCCAGCTCAGCGAAAGAGTGTCTATCAACCGGCGACCG 1386
DB 1321 ATTGCTCCAGCGAGACACAGATGCTAAAGTGAGATGTGATATCACTGACACCA 1380
QY 1387 GAAGCCCACTTCAAGGCCCAAGGACGATCTTTGGGAACTGAAAGAGAAAATCTT 1446
DB 1381 GAGGCTCAGTTCAGGCTCAGGGAAGATTTAGAAAATTAAGAAAGAAATCTTGT 1440
QY 1447 AACCCAAAGAAAGATGAAGCTGAAGGCGCATATCAGATGCGCTCTTCCACAGCTGGC 1506
DB 1441 AGTCTAAAGAAAGAGGTGAATCTTGAAGCTCATATCAGAGTGCATCTTGTGCTGGC 1500
QY 1507 CCGGTATTGCAAGAGTGGCAAGCCGTGAACGAACCTGCAAACTTAACAGTGCAGAA 1566
DB 1501 AGAGTTATTGAAAAGAGGCAAAACGTTGAATGAATCTTCAGAAATTTGTCAAGTCA 1560
QY 1567 GTATCTGCTGCTGTGACCAACGCGCATATCAGATGAGATGATGATGATGATATC 1626
DB 1561 GTTGTGTCTCTGTGTCAGACACCTGATGAGATGACCAAGTGTGTCAAAATTAAT 1620
QY 1627 GGGCACTTTCTTGTAGCCAGACTGACAGCGCAAGATCAGGAAATTGTACAAACAGT 1686
DB 1621 GGTCACTTCTATGCTTGTGAGTGGCCAGAGAAAATTCAGGAAATTCGACTCAGGTA 1680
QY 1687 AAGCAGAGAGAGAAATACCTCAGGGAGTGCTCAC 1726
DB 1681 AAGCAGACCAACAGAGAGGCTCTGCAAGTGAACAC 1720

Search completed: July 24, 2004, 07:28:53
Job time : 1231.37 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 06:34:14 ; Search time 217.862 Seconds

(without alignments)
8362.646 Million cell updates/sec

Title: US-09-270-437D-8

Perfect score: 3283

Sequence: 1 ggcagcgagagcgagcgagga.....aaccttgaaatgttattt 3283

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PTCUTS COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfilst.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3282	100.0	3283	3	US-09-061-709-8
2	3282	100.0	3283	4	US-09-899-651-8
3	3143	95.7	3412	3	US-09-061-709-6
4	3143	95.7	3412	4	US-09-899-651-6
5	697.2	21.2	1740	4	US-09-643-597-347
6	697.2	21.2	1740	4	US-09-542-615A-347
7	697.2	21.2	1740	4	US-09-606-421B-347
8	695.6	21.2	4159	3	US-09-061-709-4
9	695.6	21.2	4159	4	US-09-899-651-4
10	695.6	21.2	4181	4	US-09-643-597-175
11	695.6	21.2	4181	4	US-09-480-884A-175
12	695.6	21.2	4181	4	US-09-542-615A-175
13	695.6	21.2	4181	4	US-09-606-421B-175
14	685	20.9	2224	3	US-09-261-855-1
15	585.4	17.8	1946	3	US-09-061-709-7
16	585.4	17.8	1946	4	US-09-899-651-7
17	504	15.4	1708	3	US-09-661-709-5
18	504	15.4	1708	4	US-09-899-651-5
19	223.8	6.8	317	4	US-09-621-976-13846
20	73.8	2.2	7218	1	US-08-232-463-14
21	56.8	1.7	7218	1	US-08-232-463-14
22	45	1.4	598	4	US-09-669-751-39
23	44.6	1.4	4403765	3	US-09-103-840A-2
24	44.6	1.4	4411529	3	US-09-103-840A-1
25	44.4	1.4	5562	4	US-10-204-708-63
26	44	1.3	1664976	4	US-08-916-421B-1
27	43.4	1.3	55298	4	US-09-491-356C-1

C	28	42.8	1.3	289	3	US-09-007-005-17	Sequence 17, Appl
C	29	42.8	1.3	289	3	US-09-244-796-17	Sequence 17, Appl
C	30	42.8	1.3	17056	3	US-09-245-041-3	Sequence 3, Appl
C	31	42.2	1.3	80161	3	US-09-036-987A-1	Sequence 1, Appl
C	32	42.2	1.3	80161	3	US-09-370-700-1	Sequence 1, Appl
C	33	42.2	1.3	80161	4	US-09-603-207-1	Sequence 1, Appl
C	34	42	1.3	277	3	US-09-007-005-3	Sequence 3, Appl
C	35	42	1.3	277	3	US-09-244-796-3	Sequence 3, Appl
C	36	42	1.3	832	4	US-09-621-976-2813	Sequence 2813, Ap
C	37	41.6	1.3	248	3	US-09-007-005-32	Sequence 32, Appl
C	38	41.6	1.3	248	3	US-09-244-796-32	Sequence 32, Appl
C	39	41.4	1.3	791	3	US-08-998-416-346	Sequence 346, Appl
C	40	41.4	1.3	7130	3	US-09-056-105-31	Sequence 31, Appl
C	41	41.2	1.3	362	3	US-09-018-584A-11	Sequence 11, Appl
C	42	41.2	1.3	1722	4	US-09-134-001C-1976	Sequence 1976, Ap
C	43	41.2	1.3	56520	3	US-09-338-907-179	Sequence 179, Appl
C	44	41.2	1.3	56520	4	US-09-218-207-179	Sequence 179, Appl
C	45	41	1.2	3138	1	US-07-867-106-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-061-709-8

Sequence 8, Application US/09061709B

Patent No. 6297364

GENERAL INFORMATION:

APPLICANT: Chen, Yao-Tseng

APPLICANT: Gare, Ali

APPLICANT: Tsang, Solam

APPLICANT: Jager, Elke

APPLICANT: Knuth, Alexander

APPLICANT: Old, Lloyd J.

TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated

TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof

FILE REFERENCE: LUD 5338

CURRENT FILING DATE: 1998-04-17

NUMBER OF SEQ ID NOS: 8

SEQ ID NO 8

LENGTH: 3283

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

US-09-061-709-8

Query Match	100.0%;	Score 3282;	DB 3;	Length 3283;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3283;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	GGCAGCGAGAGCGAGAGCGCGGGGTACCGGGCGGGGAGCCCGGGGCTTCGGGG	60		
1	GGCAGCGAGAGCGAGAGCGCGGGGTACCGGGCGGGGAGCCCGGGGCTTCGGGG	60		
61	AAGAGCGATGATGAAACAAGCTTTACATCGGGAACCTGAGCCCGCGTACCGCGAC	120		
61	AAGAGCGATGATGAAACAAGCTTTACATCGGGAACCTGAGCCCGCGTACCGCGAC	120		
121	GACCTCCGAGAGCTCTTTGGGAGAGAGAGTGGCCCTGGGAGAGAGTCTCTGAAG	180		
121	GACCTCCGAGAGCTCTTTGGGAGAGAGAGTGGCCCTGGGAGAGAGTCTCTGAAG	180		
181	TCCGGCTACGCTTCGTGACTACCCGACACAGAACTGGGCGATCCGGCCATGAGAGC	240		
181	TCCGGCTACGCTTCGTGACTACCCGACACAGAACTGGGCGATCCGGCCATGAGAGC	240		
241	CTCTGGGTAAGTGAATTCATGGGAAATATCATGGAAATGTTATCTAGTCTTAA	300		
241	CTCTGGGTAAGTGAATTCATGGGAAATATCATGGAAATGTTATCTAGTCTTAA	300		
301	AAGCTAAGAGAGAGGAAATTCAGATTCGAAACATCCTCTCACTGAGTGGAGGTG	360		

```
Db 301 AAGCTAAGAGAGAGAGAAATTCAGATTGAAACATCCCTCTCACTGTGAGTGGAGGTG 360
Qy 361 TTGATGGACTTTTGGCTCAATATATGAGACATGAGAAATGTGGAAACAATCAACAGAC 420
Db 361 TTGATGGACTTTTGGCTCAATATATGAGACATGAGAAATGTGGAAACAATCAACAGAC 420
Qy 421 ACGAAACCGCGGTGTCAAGTCATATGCAACAGAGAAAGAAAGCAAAATATGACATG 480
Db 421 ACGAAACCGCGGTGTCAAGTCATATGCAACAGAGAAAGAAAGCAAAATATGACATG 480
Qy 481 GAGAAGCTAAGCGGAGATGATTTGAGAACTACTCTCTTCAGAAATTTCTTCAATCCCGAT 540
Db 481 GAGAAGCTAAGCGGAGATGATTTGAGAACTACTCTCTTCAGAAATTTCTTCAATCCCGAT 540
Qy 541 GAAAGGTGAGTCTCCCTTGCGCCCTCAAGCAGCCCAAGTGGGAGCACTTCTCCCG 600
Db 541 GAAAGGTGAGTCTCCCTTGCGCCCTCAAGCAGCCCAAGTGGGAGCACTTCTCCCG 600
Qy 601 GAGCAAGGCGACGCGCTGGGGGCACTTCTCAAGCCAGACAGATTAATTTCCGCTCGG 660
Db 601 GAGCAAGGCGACGCGCTGGGGGCACTTCTCAAGCCAGACAGATTAATTTCCGCTCGG 660
Qy 661 ATCTGTGCTCCCAACCGATTTGTGTGTCATCATCGAAAGAGGCTTGAACATTAAG 720
Db 661 ATCTGTGCTCCCAACCGATTTGTGTGTCATCATCGAAAGAGGCTTGAACATTAAG 720
Qy 721 AACATCACTAAGCAGACCCAGTCCCGGGTATGATATCCATAGAAAGAACTCTGAGCT 780
Db 721 AACATCACTAAGCAGACCCAGTCCCGGGTATGATATCCATAGAAAGAACTCTGAGCT 780
Qy 781 GAGAGAAAGCTGTACCATCCATGCGACCCAGAGGGGACTTCTGAGATGCGCATG 840
Db 781 GAGAGAAAGCTGTACCATCCATGCGACCCAGAGGGGACTTCTGAGATGCGCATG 840
Qy 841 ATTCTTGAATCATGAGAGAAAGAGCAGATGAGACCAAACTAGCCGAGAGATTCTCTG 900
Db 841 ATTCTTGAATCATGAGAGAAAGAGCAGATGAGACCAAACTAGCCGAGAGATTCTCTG 900
Qy 901 AAAATCTTGGCACAATGCGCTTGTGGAAAGCTGATTTGAAAAGAGGAGAAATTTG 960
Db 901 AAAATCTTGGCACAATGCGCTTGTGGAAAGCTGATTTGAAAAGAGGAGAAATTTG 960
Qy 961 AAAGAAATGAAATGAAACAGGAGCCAGAGTAACATCTCATCTTGCAGATTGAGC 1020
Db 961 AAAGAAATGAAATGAAACAGGAGCCAGAGTAACATCTCATCTTGCAGATTGAGC 1020
Qy 1021 AATTAACAACCCGAAAGAACCATCACTGTGAAAGGCGACAGTTGAGGCTTGCAGTCT 1080
Db 1021 AATTAACAACCCGAAAGAACCATCACTGTGAAAGGCGACAGTTGAGGCTTGCAGTCT 1080
Qy 1081 GAGATAGAGATTATGAAAGAGCTGCGTGAAGCCCTTTGAAAATGATATGCTGTGTAAC 1140
Db 1081 GAGATAGAGATTATGAAAGAGCTGCGTGAAGCCCTTTGAAAATGATATGCTGTGTAAC 1140
Qy 1141 ACCCACTCGGAAATCTTCAAGCTTGAACCCCATCAACAGTTTGGCCCGCTTCCGCA 1200
Db 1141 ACCCACTCGGAAATCTTCAAGCTTGAACCCCATCAACAGTTTGGCCCGCTTCCGCA 1200
Qy 1201 CATCACTCTTATCCAGAGAGAGATTGTGATCTTTCAATCCCAACCCAGGCTGTGGC 1260
Db 1201 CATCACTCTTATCCAGAGAGAGATTGTGATCTTTCAATCCCAACCCAGGCTGTGGC 1260
Qy 1261 GCCATCATCGGAGAGAGAGGCGGCAACATCAACAGCTGTGGAGATTGCGGAGCTCT 1320
Db 1261 GCCATCATCGGAGAGAGAGGCGGCAACATCAACAGCTGTGGAGATTGCGGAGCTCT 1320
Qy 1321 ATCAAGATGCGCTGTGGAGAGGCGCAGAGTCAAGCAAGAGAGATGTCATCATACCGG 1380
Db 1321 ATCAAGATGCGCTGTGGAGAGGCGCAGAGTCAAGCAAGAGAGATGTCATCATACCGG 1380
Qy 1381 CCAACCGGAAGCCAGTTCAGAGCCCAAGGACGATCTTTGGGAAACTGAAGAGGAAAC 1440
Db 1381 CCAACCGGAAGCCAGTTCAGAGCCCAAGGACGATCTTTGGGAAACTGAAGAGGAAAC 1440

Db 1381 CCACCGGAAGCCAGTTCAGAGCCCAAGGACGATCTTTGGGAAACTGAAGAGGAAAC 1440
Qy 1441 TTCTTTAAACCCCAAGAGAGTGAAGCTGAAAGGCAATATCAGAGTCCCTCTTCAACA 1500
Db 1441 TTCTTTAAACCCCAAGAGAGTGAAGCTGAAAGGCAATATCAGAGTCCCTCTTCAACA 1500
Qy 1501 GCTGGCCCGGATGATTTGGCAAGAGTGGCAAGACCTGTGAACGAACTGCAAACTTAACAGT 1560
Db 1501 GCTGGCCCGGATGATTTGGCAAGAGTGGCAAGACCTGTGAACGAACTGCAAACTTAACAGT 1560
Qy 1561 GCGAAGTCAATGTCCTCTGTCACCAAAACGCAAGATGAAGAAATGAGAAATGATGCTGCA 1620
Db 1561 GCGAAGTCAATGTCCTCTGTCACCAAAACGCAAGATGAAGAAATGAGAAATGATGCTGCA 1620
Qy 1621 ATTATCGGCACTTCTTTGCTAGCAGACTGCAAGCGCAAGATCAGGAAAATTTGACAA 1680
Db 1621 ATTATCGGCACTTCTTTGCTAGCAGACTGCAAGCGCAAGATCAGGAAAATTTGACAA 1680
Qy 1681 CAGGTGAAGCAGCAGAGCAGAAATACCTTCAGGAGTGTGCTTCAAGGCGAAATGTA 1740
Db 1681 CAGGTGAAGCAGCAGAGCAGAAATACCTTCAGGAGTGTGCTTCAAGGCGAAATGTA 1740
Qy 1741 GGTCTCCACAGGCAACAGCAAAACAAACGGAATGTAATGTAGCCCTTCCAAACCTGACAGAA 1800
Db 1741 GGTCTCCACAGGCAACAGCAAAACAAACGGAATGTAATGTAGCCCTTCCAAACCTGACAGAA 1800
Qy 1801 TGAGAACCAACAGCAGCAGATCGGAGCAAAACCAAGACATTTGAGAAATGAGAA 1860
Db 1801 TGAGAACCAACAGCAGCAGATCGGAGCAAAACCAAGACATTTGAGAAATGAGAA 1860
Qy 1861 GTCTGCGAGGCGGCGGAGGACTTTCGCGAGAGCTTCCGAGACCCCAAGGCGGAGAGG 1920
Db 1861 GTCTGCGAGGCGGCGGAGGACTTTCGCGAGAGCTTCCGAGACCCCAAGGCGGAGAGG 1920
Qy 1921 GCGGGGAAGTGAAGCCAGGTTTGCAGAAACCAAGCCCGCTCCCGCCCGCAGG 1980
Db 1921 GCGGGGAAGTGAAGCCAGGTTTGCAGAAACCAAGCCCGCTCCCGCCCGCAGG 1980
Qy 1981 TTCTGAGGCTTCAAGCATCACTTCACATCACTCGATCTCTCTTGAATCTCCACGA 2040
Db 1981 TTCTGAGGCTTCAAGCATCACTTCACATCACTCGATCTCTCTTGAATCTCCACGA 2040
Qy 2041 CGCTATCCCTTTTATGTTGAATTAACATAGGTGAACGTGTTCAAGCCCAAGCAAAATGCA 2100
Db 2041 CGCTATCCCTTTTATGTTGAATTAACATAGGTGAACGTGTTCAAGCCCAAGCAAAATGCA 2100
Qy 2101 ACCCTTTTCTGAGAAATGCTCTGTACATGTGTGTAATATTAAGAAAGAGATG 2160
Db 2101 ACCCTTTTCTGAGAAATGCTCTGTACATGTGTGTAATATTAAGAAAGAGATG 2160
Qy 2161 TTAAGATATGTGCTGTGGTTTACAAGGTGCTGCAAGCGGTAAATATTTTGAAT 2220
Db 2161 TTAAGATATGTGCTGTGGTTTACAAGGTGCTGCAAGCGGTAAATATTTTGAAT 2220
Qy 2221 AATATATCAATATCACTCACTCACTCAATTTTAAATTTTAAATTTTAAATTTTAAAT 2280
Db 2221 AATATATCAATATCACTCACTCACTCAATTTTAAATTTTAAATTTTAAATTTTAAAT 2280
Qy 2281 TTAAGAGAAAGAGGCTTTTCTAGACTTAAAGAAATTAAGTCTTTGGAGGTCTCAGG 2340
Db 2281 TTAAGAGAAAGAGGCTTTTCTAGACTTAAAGAAATTAAGTCTTTGGAGGTCTCAGG 2340
Qy 2341 TGTAGAGAGAGCTTTGAGGCAACCGCAAAATTTCAACCAAGAGGAAATCTGTGGA 2400
Db 2341 TGTAGAGAGAGCTTTGAGGCAACCGCAAAATTTCAACCAAGAGGAAATCTGTGGA 2400
Qy 2401 AGGACACTCAAGGAGTCTGTGATCACTGTGTATGTCAACAGAAAGGATACGCTCCT 2460
Db 2401 AGGACACTCAAGGAGTCTGTGATCACTGTGTATGTCAACAGAAAGGATACGCTCCT 2460
Qy 2461 TGAAGAGAAACTGTGCACTCCATGCGCTGTAGCTCAATACCAATTTCTTTTC 2520
Db 2461 TGAAGAGAAACTGTGCACTCCATGCGCTGTAGCTCAATACCAATTTCTTTTC 2520
```

2521 TTACAGGTTTAAATGCGTTTGTGACATACGTAATATATCTGTCCTCTGT 2580
2521 TTACAGGTTTAAATGCGTTTGTGACATACGTAATATATCTGTCCTCTGT 2580
2581 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
2581 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
2641 CCCTCATCTCAATCCGATCTAGGACACCCGCCGCCGCCGCCGCCGCCGCC 2700
2641 CCCTCATCTCAATCCGATCTAGGACACCCGCCGCCGCCGCCGCCGCCGCC 2700
2701 GTATCATCATCAACAAAGGAAAGAAAGGAAAGGAAAGGAAAGGAAAGGAA 2760
2701 GTATCATCATCAACAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 2760
2761 GGTTCCTCAAAAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 2820
2761 GGTTCCTCAAAAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGG 2820
2821 GGAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2880
2821 GGAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2880
2881 ATTTGCTTTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2940
2881 ATTTGCTTTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 2940
2941 TTTTAAAAAATGTTGAGGAAAGTAAATTTTCCAGGTTTGGCCGAGGCG 3000
2941 TTTTAAAAAATGTTGAGGAAAGTAAATTTTCCAGGTTTGGCCGAGGCG 3000
3001 AAATTCACAGTTTGTGAGGAAAGTAAATTTTCCAGGTTTGGCCGAGG 3060
3001 AAATTCACAGTTTGTGAGGAAAGTAAATTTTCCAGGTTTGGCCGAGG 3060
3061 CCTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120
3061 CCTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120
3121 TTTTAAAAAATGTTGAGGAAAGTAAATTTTCCAGGTTTGGCCGAGG 3180
3121 TTTTAAAAAATGTTGAGGAAAGTAAATTTTCCAGGTTTGGCCGAGG 3180
3181 GAGAGGTGACAGTCCAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 3240
3181 GAGAGGTGACAGTCCAGTGTGACAGTGTGACAGTGTGACAGTGTGAC 3240
3241 CCNATATGCGCTTCTTTGAGCAACCTTGAAATGTTTATTT 3283
3241 CCNATATGCGCTTCTTTGAGCAACCTTGAAATGTTTATTT 3283

RESULT 2
US-09-899-651-8
; Sequence 8, Application US/09899651
; Patent No. 6576756
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tsung
; APPLICANT: Gure, Ali
; APPLICANT: Tseng, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
; TITLE OF INVENTION: Associated
; FILE REFERENCE: LUD 5538
; CURRENT FILING DATE: US/09/899,651
; PRIORITY APPLICATION NUMBER: US/09/061,709

PRIOR FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 8
; LENGTH: 3283
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-899-651-8

Query Match 100.0%; Score 3282; DB 4; Length 3283;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCGACGGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
1 GCGACGGAG 60
61 AAG 120
61 AAG 120
121 GACCTCCGCGAGCTCTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
121 GACCTCCGCGAGCTCTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
181 TCCGCGTACGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
181 TCCGCGTACGCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
241 CTCTCGGAGTAAAGGAAATGTCATGAGAGAGAGAGAGAGAGAGAGAGAG 300
241 CTCTCGGAGTAAAGGAAATGTCATGAGAGAGAGAGAGAGAGAGAGAGAG 300
301 AAGCTAAG 360
301 AAGCTAAG 360
361 TTGATGAGAGCTTTTGGCTCAATATGAGAGAGAGAGAGAGAGAGAGAG 420
361 TTGATGAGAGCTTTTGGCTCAATATGAGAGAGAGAGAGAGAGAGAGAG 420
421 ACAGAAACCGCGCTGTCACAGTCAATATGAGAGAGAGAGAGAGAGAGAG 480
421 ACAGAAACCGCGCTGTCACAGTCAATATGAGAGAGAGAGAGAGAGAGAG 480
481 GAG 540
481 GAG 540
541 GAG 600
541 GAG 600
601 GAG 660
601 GAG 660
661 ATCTGTGCTCCCAACGAGTTTGTGATGATGATGATGATGATGATGATG 720
661 ATCTGTGCTCCCAACGAGTTTGTGATGATGATGATGATGATGATGATG 720
721 AACATCACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
721 AACATCACTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
781 GCGAG 840
781 GCGAG 840
841 ATCTTGAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
841 ATCTTGAATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

QY 901 AAAATCTTGGCAACAATGGCTTGGTGGAGAAGCTATTTGGAAAAAGGAGAAATTTG 960
Db 901 AAAATCTTGGCAACAATGGCTTGGTGGAGAAGCTATTTGGAAAAAGGAGAAATTTG 960
QY 961 AAGAAAAATTGAACATGAACAGGAGCAAGAAATACATCTCATCTTTGCAAGGATTTGAGC 1020
Db 961 AAGAAAAATTGAACATGAACAGGAGCAAGAAATACATCTCATCTTTGCAAGGATTTGAGC 1020
QY 1021 AATATCAACCCGGAAAAAGAACATCACTGTGAAGGGCAAGTTGAGGCTGTGCGAGTGT 1080
Db 1021 AATATCAACCCGGAAAAAGAACATCACTGTGAAGGGCAAGTTGAGGCTGTGCGAGTGT 1080
QY 1081 GAGATAGAGATTAAGAGAGCTGGGTGAGGCTTTGAAATGATGTGCTGTGAAC 1140
Db 1081 GAGATAGAGATTAAGAGAGCTGGGTGAGGCTTTGAAATGATGTGCTGTGAAC 1140
QY 1141 ACCCACTCCGAGTACTTCTCCAGCCTGTACCCCAATCAACAGTTTGCCCGTTCGCAAT 1200
Db 1141 ACCCACTCCGAGTACTTCTCCAGCCTGTACCCCAATCAACAGTTTGCCCGTTCGCAAT 1200
QY 1201 CATCACTCTTATCCAGAGCAAGATTTGAAATCTTCTCATCCCAACCGAGCTGTGGGC 1260
Db 1201 CATCACTCTTATCCAGAGCAAGATTTGAAATCTTCTCATCCCAACCGAGCTGTGGGC 1260
QY 1261 GGCATCATCGGGAAGAGGGGGCAACATCAACAGCTGGCGAGATTGCGCGAGCTCT 1320
Db 1261 GGCATCATCGGGAAGAGGGGGCAACATCAACAGCTGGCGAGATTGCGCGAGCTCT 1320
QY 1321 ATCAAGATTGCCCTGCGAGAGGCCCAAGCGTCAAGGAAAGATGCTCATCACCGGG 1380
Db 1321 ATCAAGATTGCCCTGCGAGAGGCCCAAGCGTCAAGGAAAGATGCTCATCACCGGG 1380
QY 1381 CCAACGGAAAGCCCAAGTTCAGAGCCCAAGGAGGAACTTTTGGGAAAATGAAAGGAAAC 1440
Db 1381 CCAACGGAAAGCCCAAGTTCAGAGCCCAAGGAGGAACTTTTGGGAAAATGAAAGGAAAC 1440
QY 1441 TTCTTTAACCCCAAGAAAGATGAAGTGAAGAGCCATATCAGATGCGCTTCTTCCACA 1500
Db 1441 TTCTTTAACCCCAAGAAAGATGAAGTGAAGAGCCATATCAGATGCGCTTCTTCCACA 1500
QY 1501 GCTGGCCGGGTGATTGGCAAAAGGTGCAAGCCGTGAACGAACTGCAAGAACTTTAACCAAGT 1560
Db 1501 GCTGGCCGGGTGATTGGCAAAAGGTGCAAGCCGTGAACGAACTGCAAGAACTTTAACCAAGT 1560
QY 1561 GCAGAACTCATCTGCTGCTGTGAACAAAGCCAGATGAAGATGAAGAGTGTGTCAGA 1620
Db 1561 GCAGAACTCATCTGCTGCTGTGAACAAAGCCAGATGAAGATGAAGAGTGTGTCAGA 1620
QY 1621 AATTATGGGGCACTTCTTGTAGCCAGACTGACAGAGCGCAAGATCAGGGAAATTTGACAA 1680
Db 1621 AATTATGGGGCACTTCTTGTAGCCAGACTGACAGAGCGCAAGATCAGGGAAATTTGACAA 1680
QY 1681 CAGGTGAAGCAGAGAGAGCAAGAAATACCTCAAGGAGTGTGCTCAACAGCCCAAGTGA 1740
Db 1681 CAGGTGAAGCAGAGAGAGCAAGAAATACCTCAAGGAGTGTGCTCAACAGCCCAAGTGA 1740
QY 1741 GGTCTCCCAAGGCAACAGCAAAAACAGATGAATGATGAGCTTCAACAGCTTGAACAGAA 1800
Db 1741 GGTCTCCCAAGGCAACAGCAAAAACAGATGAATGATGAGCTTCAACAGCTTGAACAGAA 1800
QY 1801 TGAGACCAAAACAGAGCCAGATCGGAGCAAAACCAAGAACCATTTGAGGATGAGAA 1860
Db 1801 TGAGACCAAAACAGAGCCAGATCGGAGCAAAACCAAGAACCATTTGAGGATGAGAA 1860
QY 1861 GTCTGCGAGAGCGGCAAGGAACTTGTGCGAGGCTTGAAGAACTCCAGGGGCGAGAGGG 1920
Db 1861 GTCTGCGAGAGCGGCAAGGAACTTGTGCGAGGCTTGAAGAACTCCAGGGGCGAGAGGG 1920
QY 1921 GCGGGAGAGGTCAACCCAGGTTTGGCAAGAACCAACGAGAGCCCGGCTCCGCGCCCAAGGGC 1980
Db 1921 GCGGGAGAGGTCAACCCAGGTTTGGCAAGAACCAACGAGAGCCCGGCTCCGCGCCCAAGGGC 1980
QY 1981 TTCTGAGAGGCTTGAAGCATCACTTCAACCATCCAGTGTCTCTGAACTCCACGA 2040

Db 1981 TTCTGAGAGGCTTGAAGCATCACTTCAACCATCCAGTGTCTCTGAACTCCACGA 2040
QY 2041 CGCTATCCCTTTTGTGTAACATATAGGTGAAGGTGTCAAGACCAAGCAAAATGCAAC 2100
Db 2041 CGCTATCCCTTTTGTGTAACATATAGGTGAAGGTGTCAAGACCAAGCAAAATGCAAC 2100
QY 2101 ACCCTTTTCTGTGCAAAATGCTCTGTACATGTGTGTACATATTAAGAGGAAGATG 2160
Db 2101 ACCCTTTTCTGTGCAAAATGCTCTGTACATGTGTGTACATATTAAGAGGAAGATG 2160
QY 2161 TTAAGATATGTGCTGTGGTTTACACAGGGTGTCTGACAGGGTAATTTTGAAT 2220
Db 2161 TTAAGATATGTGCTGTGGTTTACACAGGGTGTCTGACAGGGTAATTTTGAAT 2220
QY 2221 AATATATCAATTAATCAACTCACTCAATTTTATATCAATTAATTAATTTTCTT 2280
Db 2221 AATATATCAATTAATCAACTCACTCAATTTTATATCAATTAATTAATTTTCTT 2280
QY 2281 TTAAGAGAAAGCAGGCTTTTCTAGACTTTAAAGATAAAGTCTTTGGAGGTCTCACGG 2340
Db 2281 TTAAGAGAAAGCAGGCTTTTCTAGACTTTAAAGATAAAGTCTTTGGAGGTCTCACGG 2340
QY 2341 TGTAGAGAGAGCTTTGAGGCTCAACCCGCAAAATTTCAACCCGAGAGGAATCTCGCGGA 2400
Db 2341 TGTAGAGAGAGCTTTGAGGCTCAACCCGCAAAATTTCAACCCGAGAGGAATCTCGCGGA 2400
QY 2401 AGGACATCAAGGAGGCTTGTGATCACTGTGTATGTCAAGAGAAAGGATACCGTCTCT 2460
Db 2401 AGGACATCAAGGAGGCTTGTGATCACTGTGTATGTCAAGAGAAAGGATACCGTCTCT 2460
QY 2461 TGAAGAGAAACTGTGTCACTCTGTATGTCTGATCAATCACCAATTTCTCTTGGC 2520
Db 2461 TGAAGAGAAACTGTGTCACTCTGTATGTCTGATCAATCACCAATTTCTCTTGGC 2520
QY 2521 TTCAACAGGTTTAACTGTTTGGTATGATCACTGATATATATTTCTGTCTCTGTT 2580
Db 2521 TTCAACAGGTTTAACTGTTTGGTATGATCACTGATATATATTTCTGTCTCTGTT 2580
QY 2581 TATCTCTCCCTCCCTCCCTCCCTCCCTTCTCTCATCTCATTTCTTTGAATTTCTCAT 2640
Db 2581 TATCTCTCCCTCCCTCCCTCCCTCCCTTCTCTCATCTCATTTCTTTGAATTTCTCAT 2640
QY 2641 CCCTCATCTCAATCCGATATCTAGCACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCA 2700
Db 2641 CCCTCATCTCAATCCGATATCTAGCACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCA 2700
QY 2701 GTATCATATCAACAAAGGAACAAAGGAACAAAGGAACAAAGGAACAAAGGAACAA 2760
Db 2701 GTATCATATCAACAAAGGAACAAAGGAACAAAGGAACAAAGGAACAAAGGAACAA 2760
QY 2761 GGTTAATCAAAAGAAAGAGTCAATGTATGTCTAGCGTGTGGAAGAGAAACA 2820
Db 2761 GGTTAATCAAAAGAAAGAGTCAATGTATGTCTAGCGTGTGGAAGAGAAACA 2820
QY 2821 GGAACCCACCAACCAACCAATCAACCAACCAAGAAAGAAATTTCAAAAGAAAGATGT 2880
Db 2821 GGAACCCACCAACCAACCAATCAACCAACCAAGAAAGAAATTTCAAAAGAAAGATGT 2880
QY 2881 AATTGTCTTTTGTGCAATTTGTGTATGACCAATATTTCAAGCAAAATGATTTCTTTC 2940
Db 2881 AATTGTCTTTTGTGCAATTTGTGTATGACCAATATTTCAAGCAAAATGATTTCTTTC 2940
QY 2941 TTTAAAAAAAATGTGAGAGAAAGTGAATTTTCAAGGTTGTTGCGCCAGGGCGTT 3000
Db 2941 TTTAAAAAAAATGTGAGAGAAAGTGAATTTTCAAGGTTGTTGCGCCAGGGCGTT 3000
QY 3001 AAATTTCAAGATTTTAAACAGAAAAACACACAGAAAGAGTACTCAGGTCTTTTAA 3060
Db 3001 AAATTTCAAGATTTTAAACAGAAAAACACACAGAAAGAGTACTCAGGTCTTTTAA 3060
QY 3061 CTTGAGACTGTGTCTGTGTTTCTTGTGAGATTTTGAAGCTGATAGTTGAGAGCAT 3120

Db 3061 CCTCAGCACCCTGCTCTGTTCCCTTAGAGATTGTAAGCTGATAGTTGAGCAT 3120
QY 3121 TTTTATTTTATTTTATTAATAAATGAGTTGAAAAAATATAGATNCACTCCGAGCTG 3180
Db 3121 TTTTATTTTATTTTATTAATAAATGAGTTGAAAAAATATAGATNCACTCCGAGCTG 3180
QY 3181 GAGAAAGTACAGTCCAGTGTGCAACAGCTGTTGTAATGTTCTTCGCTAGCCAGAA 3240
Db 3181 GAGAAAGTACAGTCCAGTGTGCAACAGCTGTTGTAATGTTCTTCGCTAGCCAGAA 3240
QY 3241 CCNATATGCTCTTCTTTTGGAGAAACCTTGAAAAATGTTATTT 3283
Db 3241 CCNATATGCTCTTCTTTTGGAGAAACCTTGAAAAATGTTATTT 3283

RESULT 3

US-09-061-709-6
; Sequence 6, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseung
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; APPLICANT: Old, Lloyd J.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE REFERENCE: LDD 5538
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 6
; LENGTH: 3412
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-061-709-6

Query Match Best Local Similarity 95.7%; Score 3143; DB 3; Length 3412;

Matches 3283; Conservative 0; Mismatches 0; Indels 129; Gaps 1;

QY 1 GGCAAGCGAGAGCGAGAGAGCGCGGTACCGCGCGCGGAGAGCCCGGCGCTCTCGAGG 60
Db 1 GGCAAGCGAGAGCGAGAGAGCGCGGTACCGCGCGCGGAGAGCCCGGCGCTCTCGAGG 60
QY 61 AAGAGAGGATGATGAAACAGCTTTAATCGGAACTGAGCCCGCGCTGACCGCGAGC 120
Db 61 AAGAGAGGATGATGAAACAGCTTTAATCGGAACTGAGCCCGCGCTGACCGCGAGC 120
QY 121 GACCTCCGCGAGCTCTTTGGGAGACAGAAAGCTGCGCGGAGACAGGTTCTGTAAG 180
Db 121 GACCTCCGCGAGCTCTTTGGGAGACAGAAAGCTGCGCGGAGACAGGTTCTGTAAG 180
QY 181 TCCGAGTACGCTTCTGAGACTACCCGAGACAGAACTGCGGACATCCGCGCATGGAACC 240
Db 181 TCCGAGTACGCTTCTGAGACTACCCGAGACAGAACTGCGGACATCCGCGCATGGAACC 240
QY 241 CTCTCGGCTAAAGTGAATGATGAGAAATCATGGAAGTTGATTAAGTTCTTAAG 300
Db 241 CTCTCGGCTAAAGTGAATGATGAGAAATCATGGAAGTTGATTAAGTTCTTAAG 300
QY 301 AAGCTAAGAGAGAGAAATTCAGATTCGAAACATCCCTCTCACTGAGTGGAGGTG 360
Db 301 AAGCTAAGAGAGAGAAATTCAGATTCGAAACATCCCTCTCACTGAGTGGAGGTG 360
QY 361 TTGGATGAGCTTTTGGCTCAATATGGAACAGTGGAGATGGAACAGTCAACAGAC 420
Db 361 TTGGATGAGCTTTTGGCTCAATATGGAACAGTGGAGATGGAACAGTCAACAGAC 420
QY 421 ACGAAGACCGCGCTTTCAGAGTCAATATGCAACAGAGAGAGCAAAATATGCGATG 480

Db 421 ACGAAGACCGCGCTTTCAGAGTCAATATGCAACAGAGAGAGCAAAATATGCGATG 480
QY 481 GAGAAAGTACAGGCGCATCACTTGAAGAACTACTCTTCAAGATTTCTCAATCCCGAT 540
Db 481 GAGAAAGTACAGGCGCATCACTTGAAGAACTACTCTTCAAGATTTCTCAATCCCGAT 540
QY 541 GAGAAAGTACAGGCGCATCACTTGAAGAACTACTCTTCAAGATTTCTCAATCCCGAT 540
Db 541 GAGAAAGTACAGGCGCATCACTTGAAGAACTACTCTTCAAGATTTCTCAATCCCGAT 540
QY 601 GAGCAAGGCGAGCGCCCTGGGGGCACTTCTAGAGGCGAGACAGATGATTTCCCGCGG 660
Db 601 GAGCAAGGCGAGCGCCCTGGGGGCACTTCTAGAGGCGAGACAGATGATTTCCCGCGG 660
QY 661 ATCTGATCCCAACCACTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 720
Db 661 ATCTGATCCCAACCACTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 720
QY 721 AACATCACTTAAGCAGACCCAGTCCCGGTTAGATTCATAGAAAAGAACTCTGAGCT 780
Db 721 AACATCACTTAAGCAGACCCAGTCCCGGTTAGATTCATAGAAAAGAACTCTGAGCT 780
QY 781 GCAGAGAGCGCTGACCATCATGAGCCAGCCAGAGGAGACTTGAAGCATGCGCGATG 840
Db 781 GCAGAGAGCGCTGACCATCATGAGCCAGCCAGAGGAGACTTGAAGCATGCGCGATG 840
QY 841 ATCTTGAAATCATGCAAGAAAGAGAGAGATGAGCAAACTAGCCGAGAGATTCCTG 900
Db 841 ATCTTGAAATCATGCAAGAAAGAGAGAGATGAGCAAACTAGCCGAGAGATTCCTG 900
QY 901 AAAATCTTGCAACATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 960
Db 901 AAAATCTTGCAACATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 960
QY 961 AAGAAATTTGAACATGAAACAGGAGCAAGTAATCAATCTTCTGAGATTTGAGC 1020
Db 961 AAGAAATTTGAACATGAAACAGGAGCAAGTAATCAATCTTCTGAGATTTGAGC 1020
QY 1021 ATATACAAACCGGAAAGAACCATCATGTTGAAGGAGCAGTTGAGGCTGTGCGAGTCT 1080
Db 1021 ATATACAAACCGGAAAGAACCATCATGTTGAAGGAGCAGTTGAGGCTGTGCGAGTCT 1080
QY 1081 GAGATTAAGATTAAGAAAGCTGCTGAGGCTTTGAAATGATATCTGCTGTATAC 1140
Db 1081 GAGATTAAGATTAAGAAAGCTGCTGAGGCTTTGAAATGATATCTGCTGTATAC 1140
QY 1141 ----- 1140
Db 1141 ----- 1140
QY 1201 CTGTCGATGCTATGCAACAGAGGCGCCCGGAGACTCCCCCGCTGCCCCCTACAC 1260
Db 1201 CTGTCGATGCTATGCAACAGAGGCGCCCGGAGACTCCCCCGCTGCCCCCTACAC 1260
QY 1261 ----- ACCACTCTCGAATCTTCTGACGCTGTAACCCCATCAACAGTTGCGCG 1320
Db 1261 ----- ACCACTCTCGAATCTTCTGACGCTGTAACCCCATCAACAGTTGCGCG 1320
QY 1321 TTCCCGATCATCATCTTATTCAGAGAGAGAGTTGTAATCTTCTATCTCCCAACAG 1380
Db 1321 TTCCCGATCATCATCTTATTCAGAGAGAGAGTTGTAATCTTCTATCTCCCAACAG 1380
QY 1381 GCTGTGGGCGCATCATCGGAGAGAGAGGAGGAGCAGATCAAAAGCTGGGAGATTGGCC 1440
Db 1381 GCTGTGGGCGCATCATCGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
QY 1441 GAGAGCTCTATCAAGATTTGCTGCGGAGAGGCGCAGAGCTGACGCAAGAGATGATC 1500
Db 1441 GAGAGCTCTATCAAGATTTGCTGCGGAGAGGCGCAGAGCTGACGCAAGAGATGATC 1500
QY 1500 ----- 1500
Db 1500 ----- 1500
QY 1572 ATACCGGAGCAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1632
Db 1572 ATACCGGAGCAACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1632

Db 1501 ATCAACGGCCCAACGGAAAGCCAGTTCAAGGCCAGGAGCGATCTTTGGAAAGTGA 1560
 Qy 1432 GAGGAAAACTTCTTTAACCCCAAGAGAAAGTGAAGCTGGAAGCGCATATCAAGTGGCC 1491
 Db 1561 GAGGAAAACTTCTTTAACCCCAAGAGAAAGTGAAGCTGGAAGCGCATATCAAGTGGCC 1620
 Qy 1492 TCTTCCACAGCTGGCGGGGTGATGGCAAAAGGTGGCAAGACCGTGAACGAACCTGCAAG 1551
 Db 1621 TCTTCCACAGCTGGCGGGGTGATGGCAAAAGGTGGCAAGACCGTGAACGAACCTGCAAG 1680
 Qy 1552 TTAACAGTGCAGAAAGTCAATGCTGCTGCTGCAACAAAGCCAGATGAATAAGAGAAAGT 1611
 Db 1681 TTAACAGTGCAGAAAGTCAATGCTGCTGCTGCAACAAAGCCAGATGAATAAGAGAAAGT 1740
 Qy 1612 ATGCTCAGATTAATGCGGCACTTTCTTGTACCGACACTGCAACAGGCAAGATCAAGGGA 1671
 Db 1741 ATGCTCAGATTAATGCGGCACTTTCTTGTACCGACACTGCAACAGGCAAGATCAAGGGA 1800
 Qy 1672 ATTGTACAAACAGGTGAGAGAGAGAGAGAGAAATACCTCAAGGAGTGGCTTCAACAGCG 1731
 Db 1801 ATTGTACAAACAGGTGAGAGAGAGAGAGAAATACCTCAAGGAGTGGCTTCAACAGCG 1860
 Qy 1732 AGCAAGTGAAGTCTCCCAACAGGACCAACGAAACCAACGATGAATGAGCTTCCACAC 1791
 Db 1861 AGCAAGTGAAGTCTCCCAACAGGACCAACGAAACCAACGATGAATGAGCTTCCACAC 1920
 Qy 1792 CTGACAGATGAGACCAACGACGACCAAGCTGCGGAGCAAAACCAAGACCATCTGAG 1851
 Db 1921 CTGACAGATGAGACCAACGACGACCAAGCTGCGGAGCAAAACCAAGACCATCTGAG 1980
 Qy 1852 GAATGAGAGTCTGCGGAGGCGGCGAGGAGCTCTGCGAGGCGCTGAGAACCCCAAGGAG 1911
 Db 1981 GAATGAGAGTCTGCGGAGGCGGCGAGGAGCTCTGCGAGGCGCTGAGAACCCCAAGGAG 2040
 Qy 1912 CGAGAGAGGCGGAGGAGGAGTCAAGCTGAGGTTGCGAGAACCAAGGAGCTTCCCGCC 1971
 Db 2041 CGAGAGAGGCGGAGGAGGAGTCAAGCTGAGGTTGCGAGAACCAAGGAGCTTCCCGCC 2100
 Qy 1972 CCCCAGGAGCTTCTGACGAGCTTCAAGCCATCACTTCAACCATCACTGAGATCTCTGAA 2031
 Db 2101 CCCCAGGAGCTTCTGACGAGCTTCAAGCCATCACTTCAACCATCACTGAGATCTCTGAA 2160
 Qy 2032 CTCCACAGAGCTATCCCTTTAGTGAACATAAGTGAACGTTCAAGCCAGC 2091
 Db 2161 CTCCACAGAGCTATCCCTTTAGTGAACATAAGTGAACGTTCAAGCCAGC 2220
 Qy 2092 AAAATGCACACCTTTTCTGAGCAAAATGCTCTGTATCATGTGTATCAATTTGAAA 2151
 Db 2221 AAAATGCACACCTTTTCTGAGCAAAATGCTCTGTATCATGTGTATCAATTTGAAA 2280
 Qy 2152 GGGAGAGTGTAAAGATATGAGGCTTGGGTTACACAGGAGTGGCTGAGGAGTAAATAT 2211
 Db 2281 GGGAGAGTGTAAAGATATGAGGCTTGGGTTACACAGGAGTGGCTGAGGAGTAAATAT 2340
 Qy 2212 TTTAGAAATATATATCAATTAATCACTCACTCACTCAATTTTAAATCAATTAATTTT 2271
 Db 2341 TTTAGAAATATATATCAATTAATCACTCACTCACTCAATTTTAAATCAATTAATTTT 2400
 Qy 2272 TTTTCTTTTAAAGAGAAAGAGGCTTTTCTAGACTTTAAAGATTAAGTCTTTGGAG 2231
 Db 2401 TTTTCTTTTAAAGAGAAAGAGGCTTTTCTAGACTTTAAAGATTAAGTCTTTGGAG 2460
 Qy 2332 GTCTCAGAGTGTAGAGAGGCTTTGAGGCAACCCGCAAAATTCACCCAGAGGGAAT 2391
 Db 2461 GTCTCAGAGTGTAGAGAGGCTTTGAGGCAACCCGCAAAATTCACCCAGAGGGAAT 2500
 Qy 2392 CTGCTCGAAGAGACCTCAGGCGAGTCTGGATCACTGTGTATGTCAACAGAGGAGTA 2451
 Db 2521 CTGCTCGAAGAGACCTCAGGCGAGTCTGGATCACTGTGTATGTCAACAGAGGAGTA 2580
 Qy 2452 CCGTCTCTTGAAGAGGAACTCTGTCACTCTCATCTCTGTCTAGCTCATACCCATT 2511
 Db 2581 CCGTCTCTTGAAGAGGAACTCTGTCACTCTCATCTCTGTCTAGCTCATACCCATT 2640

Qy 2512 TCTCTGCTCAGAGGTTTAACTGGTTTTTGCATATGCTATATATCTCTGTCT 2571
 Db 2641 TCTCTGCTCAGAGGTTTAACTGGTTTTTGCATATGCTATATATCTCTGTCT 2700
 Qy 2572 CTCTCTGTTATCTCTCCCTCCCTCCCTCCCTCTTCTCTCATCTCATCTTTTGA 2631
 Db 2701 CTCTCTGTTATCTCTCCCTCCCTCCCTCCCTCTTCTCTCATCTCATCTTTTGA 2760
 Qy 2632 TTTTCTCANTCCCTCATCTCAATCCCGTATCTACGCAACCCCTCCCTCCCAAGCA 2691
 Db 2761 TTTTCTCANTCCCTCATCTCAATCCCGTATCTACGCAACCCCTCCCTCCCAAGCA 2820
 Qy 2692 GTGCTGAGTATCATCATCAACAAAGAGCAAAAGCGAAACACACAAACGCTCAA 2751
 Db 2821 GTGCTGAGTATCATCATCAACAAAGAGCAAAAGCGAAACACACAAACGCTCAA 2880
 Qy 2752 CTTCACCTGTTTACTCAAAAGACAAAGTCAATGTCTGTCTGCTGCTTTGGA 2811
 Db 2881 CTTCACCTGTTTACTCAAAAGACAAAGTCAATGTCTGTCTGCTGCTTTGGA 2940
 Qy 2812 AGGAAACAGAAACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAATG 2871
 Db 2941 AGGAAACAGAAACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAATG 3000
 Qy 2872 AAAGATGATTTTGTCTTTTGTGATTTGTGTATGAGCCATCATATTCAAGAAATG 2931
 Db 3001 AAAGATGATTTTGTCTTTTGTGATTTGTGTATGAGCCATCATATTCAAGAAATG 3060
 Qy 2932 ATTCTCTTTTAAAAAATGAGAGAAAGTGAAGAAATTTACAGAGTGTGGCC 2991
 Db 3061 ATTCTCTTTTAAAAAATGAGAGAAAGTGAAGAAATTTACAGAGTGTGGCC 3120
 Qy 2992 CAGGAGCTTAAATTCACAGATTTTAAAGAGAAACACACAGAGAGTCACTCAG 3051
 Db 3121 CAGGAGCTTAAATTCACAGATTTTAAAGAGAAACACACAGAGAGTCACTCAG 3180
 Qy 3052 GTGTTTTTACCTCAGACCTTGTCTTGTGTTTCCCTTGAAGATTTTGAAGTATG 3111
 Db 3181 GTGTTTTTACCTCAGACCTTGTCTTGTGTTTCCCTTGAAGATTTTGAAGTATG 3240
 Qy 3112 TTGAGACATTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3171
 Db 3241 TTGAGACATTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300
 Qy 3172 GCCAGCTGAGAGAGTGAAGTCAAGTCAAGTGTGCAAGCTGTCTGAATGTCTCGCT 3231
 Db 3301 GCCAGCTGAGAGAGTGAAGTCAAGTCAAGTGTGCAAGCTGTCTGAATGTCTCGCT 3360
 Qy 3232 AGCCAAAGACNATATGCTTCTTTTGAACAACTTGAAATGTTTATTT 3283
 Db 3361 AGCCAAAGACNATATGCTTCTTTTGAACAACTTGAAATGTTTATTT 3412

RESULT 4
 US-09-899-651-6
 ; Sequence 6, Application US/09899651
 ; Patent No. 6576756
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Yao-Tseng
 ; APPLICANT: Gare, Ali
 ; APPLICANT: Tsang, Solam
 ; APPLICANT: Stockert, Elisabeth
 ; APPLICANT: Jager, Elke
 ; APPLICANT: Kuntz, Alexander
 ; APPLICANT: Old, Lloyd J.
 ; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
 ; TITLE OF INVENTION: Associated
 ; FILE REFERENCE: LUD 5538
 ; CURRENT APPLICATION NUMBER: US/09/899,651
 ; CURRENT FILING DATE: 2001-07-06
 ; PRIORITY APPLICATION NUMBER: US/09/061,709


```
Db 1981 GATTAGAAAGTCTGCGAGAGCGGCGGAGCACTCTCCAGGCGCTGAGAAACCCGAGG6C 2040
Qy 1912 CGAGAGAGGCGCGGAGAGGTGACGCCAGGTTTCCAGAACCAACGAGCCCGCTCCGCGC 1971
Db 2041 CGAGAGAGGCGCGGAGAGGTGACGCCAGGTTTCCAGAACCAACGAGCCCGCTCCGCGC 2100
Qy 1972 CCGCAGGCGCTTGTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 2031
Db 2101 CCGCAGGCGCTTGTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 2160
Qy 2032 CCGCAGGCGCTTGTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 2091
Db 2161 CCGCAGGCGCTTGTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGG 2220
Qy 2092 AAAATGACACCCCTTTTCTGTGGCAATGCTGTGTGATGATGATGATGATGATGATGAT 2151
Db 2221 AAAATGACACCCCTTTTCTGTGGCAATGCTGTGTGATGATGATGATGATGATGATGAT 2280
Qy 2152 GGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2211
Db 2281 GGGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Qy 2212 TTAGAAATATATATATATATATATATATATATATATATATATATATATATATATAT 2271
Db 2341 TTAGAAATATATATATATATATATATATATATATATATATATATATATATATATAT 2400
Qy 2272 TTTTCTTTTAAAGAGAGAGGCTTTTCTAGACTTTTAAAGATTAAGTCTTTGGGAG 2331
Db 2401 TTTTCTTTTAAAGAGAGAGGCTTTTCTAGACTTTTAAAGATTAAGTCTTTGGGAG 2460
Qy 2332 GCTCTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2391
Db 2461 GCTCTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Qy 2392 CTCTGCGAAGGACACTGACGCGAGTTCTGATCACTGTGATGATGATGATGATGATGAT 2451
Db 2521 CTCTGCGAAGGACACTGACGCGAGTTCTGATCACTGTGATGATGATGATGATGATGAT 2580
Qy 2452 CCGCTCTCTTGAAGAGAGAACTCTGTCACTCTCTCATGCTGTCTAGCTTACCACTT 2511
Db 2581 CCGCTCTCTTGAAGAGAGAACTCTGTCACTCTCTCATGCTGTCTAGCTTACCACTT 2640
Qy 2512 TCTCTTGTCTGACAGGTTTAACTGTTTGAATGATGATGATGATGATGATGATGATGAT 2571
Db 2641 TCTCTTGTCTGACAGGTTTAACTGTTTGAATGATGATGATGATGATGATGATGATGAT 2700
Qy 2572 CTCTCTGTTATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2631
Db 2701 CTCTCTGTTATCTCTCCCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCT 2760
Qy 2632 TTCTCTCATCTCTCATCTCTCATCTCTCATCTCTCATCTCTCATCTCTCATCTCTCAT 2691
Db 2761 TTCTCTCATCTCTCATCTCTCATCTCTCATCTCTCATCTCTCATCTCTCATCTCTCAT 2820
Qy 2692 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2751
Db 2821 GTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880
Qy 2752 CTTCACCTTGTATCTCAAAAGAACAGATGATGATGATGATGATGATGATGATGATGAT 2811
Db 2881 CTTCACCTTGTATCTCAAAAGAACAGATGATGATGATGATGATGATGATGATGATGAT 2940
Qy 2812 AGGAAACAGAGAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 2871
Db 2941 AGGAAACAGAGAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 3000
Qy 2872 AAAGAAATGATTTTGTCTTTTGTGATTTGTGATTTGTGATTTGTGATTTGTGATTTGT 2931
Db 3001 AAAGAAATGATTTTGTCTTTTGTGATTTGTGATTTGTGATTTGTGATTTGTGATTTGT 3060
Qy 2932 ATTCTTTCTTTAAAAAAAATGATGAGAGAGATGATGATGATGATGATGATGATGATGAT 2991
```

```
Db 3061 ATTCTTTCTTTAAAAAAAATGATGAGAGAGATGATGATGATGATGATGATGATGATGAT 3120
Qy 2992 CAGGCGCTTAATTCAGAGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3051
Db 3121 CAGGCGCTTAATTCAGAGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
Qy 3052 GTGTTTAACTCAGACCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3111
Db 3181 GTGTTTAACTCAGACCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
Qy 3112 TTGAGCATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 3171
Db 3241 TTGAGCATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 3300
Qy 3172 GCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3231
Db 3301 GCCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
Qy 3232 AGCCAAAGAACCAATATGCTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3283
Db 3361 AGCCAAAGAACCAATATGCTTTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3412

RESULT 5
US-09-643-597-347
; Sequence 347, Application us/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongrong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-643-597-347

Query Match 21.2%; Score 697.2; DB 4; Length 1740;
Best Local Similarity 64.9%; Pred. No. 4.8e-185;
Matches 1116; Conservative 0; Mismatches 538; Indels 66; Gaps 3;
```

Db 241 CGGAAGCTTCAGATGCGAATATCCCGCTCATTTACAGTGGAGGTGCTGATGTTTA 300
 QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGGAACAGTGAACAGACAGAAACCGCC 432
 Db 301 CTAGTCCAGTATGAGTGTGGAGAGCTGTGACAGTGAACAATGACCTGGAAACCTGA 360
 QY 433 GTTGTCAACGTACATATGCAACAAGAGAGAGAGAAATATGACATGGAGAGTAAAC 492
 Db 361 GTTGTAAATGTAACCTATTCAGTATAGACCAAGCTTACAGACACTGACAACTGAT 420
 QY 433 GGGCATCAGTTTGAAGAACTACTCTCTTCAAGATTTCTTCAATCCCGATGAGAGGTGAGC 552
 Db 421 GGATTTCAAGTATGAGAAATTTCACTTGAAGTAAAGCTTATATCCCTGATGAAACGGCCGCC 480
 QY 553 TCCCCCTTGCCCCCTCAGCGAGCCAGAGGTGGGGACCACTCTTCCCGGAGCAAGGC--- 609
 Db 481 CAGCAAAACCCCTTGACAGAGCCCGGAGGTGCGCGGGGCTTGGGCAAGGGGCTCTCA 540
 QY 610 -----CAGGCCCTGGGGGACACTTCTCAGGCGACAGATTTGATTTCCCGGTGGATC 663
 Db 541 AGGAGGGGTCTCCAGAGATCCGTATCCAGAGAGAAACATGATTTGCTCTGCGCTG 600
 QY 664 CTGGTCCCAACCCAGTTTGTGTGTCATTCGAAAGAGGGCTTGACCATTAAGAAC 723
 Db 601 CTGGTCCCAACCAATTTGTGAGCCATCATAGGAAAGAGGTGCGACCATTCGAGAC 660
 QY 724 ATCACTAGCAGACCCAGTCCCGGGTAGATTCATAGAAAGAGAACTCTGAGCTGCA 783
 Db 661 ATCCCAACAGACCCAGTCTTAATAATCATGTTCACCGTAAAGAAATGCGGGGCTGCT 720
 QY 784 GAGAGGCTGTCACTTCATCCAGCCAGAGGGGACCTTGTGAAGCATGCGCATAT 843
 Db 721 GAGAGGCTGTCTATCTCTCTCTATCTCTGAGAGCACTTGGCGCTTGAATCTAT 780
 QY 844 CTGTAATCATGCAAGAGAGGAGAGTGAAGCAAACTAGCGGAGATTCCTGTA 903
 Db 781 CTGAGATATATGCTAAGAGAGCTCAAGATTAATAATTCAGAGAGATCCCTTGAG 840
 QY 904 ATCTGGCACAATGCTGTGTGTAAGAGCTGATTTGAAAGAGAGCAAAATTTGAG 963
 Db 841 ATTTTACCTATATTAATCTTGTGTGAGCTTTATTTGTATTAAGAGAGAAATCTTAA 900
 QY 964 AAAATGACATGAAACAGGAGCCAGATTAACAATCTCATCTTGTGAGATTTGAGATA 1023
 Db 901 AAAATGAG 960
 QY 1024 TACAAACCGAG 1083
 Db 961 TATATATCAG 1020
 QY 1084 ATGAGAGATTAAG 1143
 Db 1021 GAGAGAGATTAAG 1080
 QY 1144 CACTCCGAT-----ACTTCCCA 1162
 Db 1081 CAGAGCATTTAATCTCGATTAATAATCTGAAGCCTTGAGTCTGTTCCACCACTTCA 1140
 QY 1163 GCTGTAAACCCCATCACAGTTTGGCCGTTCCCGCATCACTC----- 1208
 Db 1141 GGGATGCACTCCCATCTCAGAGGCCCCCTTCAAGCCATGACTCTCCCTACCCGAGTTT 1200
 QY 1209 --TTATCCAGAGAGAGAGATTTGATCTCTTATCTCCCAACCAAGGTGTGGGGCCATC 1266
 Db 1201 GAGCAATCGAAG 1260
 QY 1267 ATCGGAG 1326
 Db 1261 ATCGGAG 1320
 QY 1327 ATTGCCCCCTGGAG 1386

Db 1321 ATTGCTCAGGAG 1380
 QY 1387 GAGGCCAGTTTCAAGGCCAG 1446
 Db 1381 GAGGCTCAGTTTCAAGGCCAG 1440
 QY 1447 AACCCAAAG 1506
 Db 1441 AGCTTAAAG 1500
 QY 1507 CGGGATTTGGCAAG 1566
 Db 1501 AGAGTTATGAG 1560
 QY 1567 GTCATGTGCTCTGTGACCAAG 1626
 Db 1561 GTTGTGTCTCTGTGACCAAG 1620
 QY 1627 GGGGACTTTTGTGAG 1686
 Db 1621 GGTCACTTATGCTTGTGAG 1680
 QY 1687 AAGCAG 1726
 Db 1681 AAGCAG 1720

RESULT 6
 US-09-542-615A-347
 ; Sequence 347, Application US/09542615A
 ; Patent No. 6518256
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Hoeken, Nancy A.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; FILE REFERENCE: AND DIAGNOSIS OF LUNG CANCER
 ; CURRENT APPLICATION NUMBER: US/09/542,615A
 ; CURRENT FILING DATE: 2000-04-14
 ; NUMBER OF SEQ ID NOS: 350
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 347
 ; LENGTH: 1740
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-542-615A-347

Query Match 21.2%; Score 697.2; DB 4; Length 1740;
 Best Local Similarity 64.9%; Pred. No. 4.8e-185;
 Matches 1116; Conservative 0; Mismatches 538; Indels 66; Gaps 3;

QY 73 ATGACAAAGCTTTACATCGGAAACCTGAGCCCGCGTCAACCGCCGAGCTCCGCGAG 132
 Db 1 ATGACAAAGCTTTACATCGGAAACCTGAGCCCGCGTCAACCGCCGAGCTCCGCGAG 60
 QY 133 CTCTTGGGAGACAGAGAGCTGCCCTGGGGGAGCAGGTCTCTGCTGAAGTCCGGGTAGCC 192
 Db 61 ATCTTCAAGAGAGCCCAAGATCCGGGTGTGGAGACCTCTCTGTGAGAGACTGGGTAGCG 120
 QY 193 TTGTGAGACTACCCGAG 252
 Db 121 TTGTGAG 180
 QY 253 GTGAGATTTGATGAG 312
 Db 181 ATGAG 240
 QY 313 AGGAAATTCAGATTGAAACATCCCTCTCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372

```

Db      241 CGGAATCTTGAGATACGAAATATCCCGCTCATTTACAGTGGAGGTGTGATAGTTTA 300
Qy      373 TTGGCTCAATATGGGACAGTGAAGATGTGAACAAGTCAACACAGACACAGAAACCGCC 432
Db      301 CTAGTCCAGTATGAGTGTGTGAGAGCTGTGAGCAAGTGAACACTGACTCGGAAACTGCA 360
Qy      433 GTTGTCAACGTCACTATGCAACAGAGAAAGAAATATGCGCATGAGAGATGAG 492
Db      361 GTTGTAAATGTAACTTATTCAGTAGAGCCAGCTTACAGACAGACTGACAACTGAT 420
Qy      493 GGGCATCAGTTTGAAGACTACTCTTCAAGATTTCTCATCTCCCGATGAAAGGTGAGC 552
Db      421 GGATTTCACTTGAAGATTTTCACTTGAAGTACCTTATCTCCGATGAAACCGCCGCC 480
Qy      553 TCCCTTGGCCCTTCAAGCAGCCAGCGTGGGAGCACTCTTCCCGGAGCAAGC--- 609
Db      481 CAGCAAAACCCCTTCAAGCAGCCAGCGTGGGAGCACTCTTCCCGGAGCAAGC--- 540
Qy      610 -----CAGCCCTTGGGAGCACTTCTAGGCGCAGACAGATTGTTCCCGCTGCGGATC 663
Db      541 AGGCAAGGAGTTCAGAGATCGTATCCAGAGAAACCATGTGATTTGCTTGGCGCTG 600
Qy      664 CTGCTCCCACTTGTGTTGTGTCATCATCGAAAGAGGCGCTTGAACATAAAAGAC 723
Db      601 CTGCTTCCCACTTGTGTTGTGTCATCATAGGAAAGAGGCGCACTTGGGAGC 660
Qy      724 ATCTAAGCAGACCCAGTCCCGGAGATATCCATAGAAAGAAAGAACTTGGAGCTGCA 783
Db      661 ATCACCAGAACAGACCCAGTCTAAATTCATGTCACCGTAAAGAAATCGGAGGCTGT 720
Qy      784 GAGAGCCGTGACCATGATGATGCAACCCAGAGGAGCTTGTGAAGCATGCGCATGAT 843
Db      721 GAGAGTGCATATCTATCTCTTCTTACTCTTGAAGGACCTTGGCGCTTGTAAAGTAT 780
Qy      844 CTTGAATCATGACAGAAAGAGCAGATGAGACCAATGACCGAAGATTCCTCTGAAA 903
Db      781 CTGGAGATATGATAGAGAGAGCTCAAGATATTAATTCACAGAAAGATCCCTTGAAG 840
Qy      904 ATCTTGGCACACATGCGCTTGTGGAAGACTGATTTGAAAGAAAGAGGAAATTTGAAG 963
Db      841 ATTTTATGCTCATATATCTTGTGTGAGCGTCTTATTTGTAAGAAAGAAATCTTAAA 900
Qy      964 AAAATGACATTAACAGAGGACAGATTAACATCTCATCTTGTGAAGATTTGACATA 1023
Db      901 AAAATGACAGACACAGACACTTAAATCAGATATCTCATTTGAGAGATTTGAAGCTG 960
Qy      1024 TACAACTCCGAGAAAGAACATCACTGTGAAGGAGCAGTTGAGGCTGTGCACTGTGAG 1083
Db      961 TATTAATCCAGAAAGCACTATTAAGTTAAAGCAATGTTGAGACATGTGCCAAAGCTGAG 1020
Qy      1084 ATAGAGATTAAGAAAGCTGCGTGAAGCTTTGAAATATGATGCTGCTGTTAAACACC 1143
Db      1021 GAGGAGATCATGAAAGAAATCAGGAGCTTATGAAATATGATGCTTAAAGATCTT 1080
Qy      1144 CACTCCGAGAT-----ACTTCTCCA 1162
Db      1081 CAGGCACTTAAATCTCTGATTAATCTGAAGCGCTTGGGTCTGTTCCACCACTTCA 1140
Qy      1163 GCGCTGATACCCCATCAACAGTGTGGCCGCTTCCCGCATCATCTC----- 1208
Db      1141 GGGATGCACTCTCCACTCAAGGAGCCCTTCAAGCCATGACTCTCTCCCTTCAACCGCAGTT 1200
Qy      1209 --TTATTCAGAGAGAGATTTGTAATCTCTCATCCCAACCGAGGTGTGGGCGCATC 1266
Db      1201 GAGCAATCAAGAAAGGAGACTGTTCATCTGTTATCCAGCTCTATATAGTCGGTCAATC 1260
Qy      1267 ATCGGAGAAAGAGGAGCAGACATCAAAACAGCTGCGAGATTTCCCGGAGCCTTATCAAG 1326
Db      1261 ATGGGCAAGAGAGGCGCAGACATCAAGACACTTCTGCGTTGTGAGGCTTCAATTAAG 1320
Qy      1327 ATTGCCCTTCCGAAAGGCCAGACGTGAGGAAAGATGTCATCATCACCGGCGCACCG 1386
Db      1321 ATTGCTCCAGCGGAGAGCAGACAGATGTAAAGTGAATGTGATATCATGAGCACCA 1380

```

```

Qy      1387 GAAGCCAGTTCAAGGCCACAGGACGATCTTTGGGAACTGAAAAGAGAAACTTCTTT 1446
Db      1381 GAGGCTCAGTTCAAGGCTCAGAGAAAGATTTATGAAAAATTAAGAAAGAAACTTGT 1440
Qy      1447 AACCCAAAGAAAGTGAAGCTGGAAGCCGATATCAAGTGGCCCTTCCACAGCTGGC 1506
Db      1441 AGCTTAAAGAAAGAGTGAAGCTTGAAGCTCATATCAAGTGGCCATCTTGTGCTG 1500
Qy      1507 CGGATGATTTGGCAAGGTGGCAAGACGTGAACGAACTGCAAGAACTTAACAGTGCAGAA 1566
Db      1501 AGAGTTATGAAAGAGGAGGAGAAAGGTGAATGAATCTTCAAGTTGTCAAGTGCAGAA 1560
Qy      1567 GTCATCTGCTCTGTGACCAACCGCAGATGAAATTAAGAAAGTATGCTAAGAAATTATC 1626
Db      1561 GTTGTGTCCCTCGTACGACAGACCTGATGAGATGACAAAGTGTGTGCAAAATTA 1620
Qy      1627 GGGCACTTCTTGTGAGCAGACTGCAACAGCGCAAGATCAGGAAATTTGTAACAGGTG 1686
Db      1621 GGTCACTTCTATGCTTGTGCAAGTGTGCCAGAGAAATTCAGGAAATTTCTGACTGAGTA 1680
Qy      1687 AAGCAGAGAGACAGAAATACCTCAGGAGTCCGCTGAC 1726
Db      1681 AAGCAGACCAACACAGAGGCTTGCAGAAAGTGAACAC 1720

```

```

RESULT 7
US-09-606-421B-347
; Sequence 347, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Skelky, Yasir A.W.
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210.121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-421B-347

```

```

Query Match      21.2%; Score 697.2; DB 4; Length 1740;
Best Local Similarity 64.9%; Pred. No. 4.8e-185;
Matches 1116; Conservative 0; Mismatches 538; Indels 66; Gaps 3;

Qy      73 ATGAACAGCTTATACATCGGAACTGAGCCCGCGCTCAACGCGCAGACGACTTCGCGAG 132
Db      1 ATGAACAAATCTGATATCGGAACTCAAGAGAGCGCGCCCTTCGAGCTTGAAGAT 60
Qy      133 CTCTTTGGGAGAGAACTGCGCTTGGCGGAGAGTCTCTGTGAAGTCCGGCTACGCC 192
Db      61 ATCTTCAAGAGACCCCAAGATCCCGGTGCGGACCTTCTCTGGTGAAGACTGGCTACGCG 120
Qy      193 TTGGTGAACATACCCGACAGAACTGGGCGATCCGCGCATTCAGACCTCTCGGATAA 252
Db      121 TTGTGAGACTGCGCGAGCGAGAGTGGCCCTCAAGCCATCGAGGCGCTTTGAGTAAA 180
Qy      253 GTGGAATTCATGGGAAATCATGGAATTGATTACTCACTCTCTTAAAGCTTAAAGAGC 312
Db      181 ATGAATCTGACGCGGAAACCATGAAAGTTGAGCATCTGGTCCCAAAAGCAAGAT 240

```

```

QY 313 AGAATAATTCAGATTGGAACATCCCTCCTCACTGACGTGGAGGTGTGGATGACTT 372
DB 241 CGAATACTCAGATAGAGAAATATCCGCTCTATTTACATGGAGAGTGTGATGATT 300
QY 373 TTGGCTCATATGGACAGATGGAATGTGAACAAGTCAACACAGACACAGAAACCGCC 432
DB 301 CTAGTCCAGATATGGATGTGTGAGCTGTGAGCAAGTGAACATGACCTCGAACTGCA 360
QY 433 GTTGTCAAGTCACATATGCAACAAGAGAAAGAAATAATGCCATGGAGAAAGTAAAGC 492
DB 361 GTTGTAAATGTAACTATTCCTAGTAAAGACCAAGCTAGCAAGCACTAGCAAACTGAT 420
QY 493 GGGCATCAAGTTGAGAACTACTCTTCAAGATTTCTCATATCCCGATGAAGAGTGAAC 552
DB 421 GGATTTCAAGTTAGAAATTTCACTTGAAGAGAGCTATATTCCTGATGAAGACGCCGCC 480
QY 553 TCCCTTGGCCCCCTCAGCGAGCCAGAGGTGGGACCACTCTTCCGGGAGCAAGC--- 609
DB 481 CAGCAAAACCCCTTGCACAGAGCCGAGGTGCGCCGGGGCTTGGGCAAGGGGCTCTCA 540
QY 610 -----CAGGCCCCCTGGGGGCACTTCTCAGGCGCAGACAGATTGATTTCCCGCTGGATC 663
DB 541 AGGAGAGGGGTCTCAGAGATCCGTATTCAGAGAAACCATGTGATTTGCTCTGGGCTG 600
QY 664 CTGTATCCCAACCAAGTTTGTGTGTCATCATCGAAAGAGAGGCTTGACCATAAAGAC 723
DB 601 CTGTATCCCAACCAATTTGTTGAGCCATCATAGSAAAGAGGCCACATTCGGAGAC 660
QY 724 ATCACTAAGCAGACCCAGTCCCGGGTGAATATCCATAGAAAGAACTTGGAGCTGCA 783
DB 661 ATCACTAAGCAGACCCAGTCCCGGGTGAATATCCATAGAAAGAACTTGGAGCTGCA 720
QY 784 GAGAGAGCTGTACCATCATGACCAACCCAGAGGGAGATTGTGAAGCATGCCGATGAT 843
DB 721 GAGAGAGCTGTATCTATCTCTTCTTCTTAAAGGACACTCTGCGCTTGAAGTCTAT 780
QY 844 CTTGAAATCATGACAGAAAGAGCAGATGAGACCAACTGAGCGAAGATCTCTGAAA 903
DB 781 CTGAGAGATTATGCTAAGAAAGCTCAGATATTAATTCACAGAGAGATCCCTTGAAG 840
QY 904 ATCTGGCACAATATGCTGTGTGGAAGATGAGATGTTGAAAAGAGCAGAAATTTGAG 963
DB 841 ATTTTACTCTATATTAATTTGTTGAGCTTATGTTGTTAAAGAGAAATCTTAA 900
QY 964 AAAATTGACATGAAACAGGGACCAAGATTAATATCTCATCTTTCAGAGATTTAGATA 1023
DB 901 AAAATTGACATGAAACAGAGACCTAAATATCAGATATCTCATTTGAGAAATGACCTG 960
QY 1024 TACAACCCGGAAGAACATCATCTGTAGAGGGCAGATTAGGCTGTGACAGTGTAG 1083
DB 961 TATATATCAGAGACATATTAAGTTAAAGGCAATTTGAGCATGTGCCAAAGCTGAG 1020
QY 1084 ATGAGATTTTGAAGAACTGCTGAGGCTTTGAAAATGATATGCTGCTTTAAACCC 1143
DB 1021 GAGAGAGTATGAGAAATCAGGGAGTCTTATGAAATATGTTGCTTATGAAATCTT 1080
QY 1144 CACTCCGAGT-----ACTTCTCCA 1162
DB 1081 CAGAGCATTTAATCTCTGATTAATCTGAAGCTTGGGTCTGTTCCACCACTTCA 1140
QY 1163 GCGCTGACCCCATCAAGATTGTGCCCCCTTCCGATCATCACTC----- 1208
DB 1141 GGGATGCACTCTCCACCTCAGGGCCCCCTTTCAGCCATGACTCTCTCCATACCCGAGTTT 1200
QY 1209 ---TTATCCAGAGAGAGATTTGAATCTCTTATCCCAACCCAGAGCTGTGGGCGCATC 1266
DB 1201 GAGCAATTCAGAAACGAGACTGTTCACTGTTATATCCAGCTCTATCAGTCCGTGCATC 1260
QY 1267 ATCGGAGAAAGAGGGGACACATCAAAAGCTGGCGAGATTCGCCGAGCTCTATCAAG 1326
DB 1261 ATCGGAGAGAGGGGACACATCAAAAGCTGTTCTGCTTGTGAGAGCTTCAATTAAG 1320
QY 1327 ATTGCCCCCTGCGGAGGCCAGACGTACGAGAAAGATGTCATCATCACCGGGCCACCG 1386

```

```

DB 1321 ATGTCTCAGCGGAGACACCATCTAAAGTGAAGATGTGATATCATGTGACACCA 1380
QY 1387 GAAGCCAGTTCAAGGCCCAGGAGCGAGTCTTTGGGAAATGAAAGAGAAATCTCTTT 1446
DB 1381 GAGGCTCAGTTCAAGGCTCAGGAGAAATTTATGAAAAATTTAAAGAAATCTTTGTT 1440
QY 1447 AACCCAAAGAAAGATGAGAGCTGAGAGCGCATATCAGAGTCCCTCTTCCACAGTGGC 1506
DB 1441 AGTCTTAAAGAAAGATGAGAAATCTTAAGCTCATATCAGAGTCCATCTTGTCTGTGCG 1500
QY 1507 CGGGTATTTGCAAGAGTGGCAAGACCTGAAACGAACTGCAAGACTTAAACAGTCAGAA 1566
DB 1501 AGAGTTATGAAAAAGAGGAGCAAAACGATGATGAATCACTCAGAAATTTGTCAAGTCAGAA 1560
QY 1567 GTCATCTGCTCGTGCAGCAACCAAGCCAGATGAAATGAGGAAGTATGTCAGAAATTATC 1626
DB 1561 GTTGTGTCTCTCTGTGACCAAGACCTATATGAGATGACAAAGTGTGTCAAAATTAAT 1620
QY 1627 GGGCACTTCTTGTCTAGCCAGATGACAGGCGCAAGATCAGGAAATTTGTAACAAGGTG 1686
DB 1621 GGTCACTTCTATGTCTTGGCCAGGTGGCCAGAGAAATTTCAAGAAATTTCTGACTCAGGTA 1680
QY 1687 AAGCAGAGAGCAGAAATATCCCTCAGGAGTCCCTTAC 1726
DB 1681 AAGCAGAGCAGCAACAGAAAGCTCTGCAAGTGGACAC 1720

RESULT 8
us-09-061-709-4
; Sequence 4, Application US/09061709B
; Patent No. 6297364
; GENERAL INFORMATION:
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Gure, Ali
; APPLICANT: Tsang, Solam
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Jager, Elke
; APPLICANT: Knuth, Alexander
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
; FILE REFERENCE: LUD 5538
; CURRENT APPLICATION NUMBER: US/09/061,709B
; CURRENT FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 8
; SEQ ID NO 4
; LENGTH: 4159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
;
us-09-061-709-4

Query Match      21.2%; Score 695.6; DB 3; Length 4159;
Best Local Similarity 64.8%; Pred. No. 2.5e-184;
Matches 1115; Conservative 0; Mismatches 539; Indels 66; Gaps 3;

```

QY	313	AGGAAATTTAGATTGGAAACATCCCTCTCACCTGCACTGGAGAGTGTGGATGCACTT	372
Db	491	CGGAAATCTTAGATATACAAATATATCCGCTCATTTACAGTGGAGAGTGTGATAGTTTA	550
QY	373	TTGGCTCAATATGGGACAGTGGAGAAATGTGGAAACAAGTACAACAGACAAGAAACCGCC	432
Db	551	CTAGTCCAGATATGGAGATGTGTGGAGAGTGTGAGACATGTGAAACACTGACTCGGAAACTGCA	610
QY	433	GTTGTCAACGTACATATGCAACAGAGAAAGAAACAAATATAGCATGTGAGAAAGCTAAGC	492
Db	611	GTTGTAATATGTAACCTATTCAGTAGAGCAAGACTAGCAACACTTAACAAACCTGAAT	670
QY	493	GGGCAATCAGTTTGGAACTACTCCCTTCAAGATTTCCATACCTCCGGATGAAGGTGAGC	552
Db	671	GGATTTCAGTTAGAAATTTCACTTTAAGATGACCTATATCCCTGATGAATGGCCGCC	730
QY	553	TCCCTTCGCCCCCTCAGCGAGCCAGCGTGTGGGAGCCACTCTTCCGGAGCAAGGC---	609
Db	731	CAGCAAAACCCCTTGCAAGACCCCGAGAGTGGCCGGGGCGCTTGAGCAAGAGGAGCTCTCA	790
QY	610	-----CAGCGCCCTGGGGGCACTTCTGAGGCCAGACATTTGATTTCCGCTGCGGATC	663
Db	791	AGCGAGGGGCTCCAGAGATCCGATATCCAGAGAGAAACATGATATTTGCTGCGCCTG	850
QY	664	CTGTGTCCTCCACCCAGTTGTTGTGGCATCATATCGGAAAGAGGGGCTTGACATTAAGAAC	723
Db	851	CTGTGTTCCACCCCAATTTGTGGAGCCATCATAGAAAGAAAGAGGTGCACATTGCGAAC	910
QY	724	ATCACTAAGCAGAACCCAGTCCCGGGATGATATCCATAGAAAGAGAACTGTGAGCTGCA	783
Db	911	ATCACCAAACAGACCCAGCTTAATAATGATGTTCACCCGTAAGAAATATGGGGGGCTGCT	970
QY	784	GAGAAAGCTGTCAACCATCATGATCCACCCAGAGGGGACTTTGAAAGATGCCAGATGAT	843
Db	971	GAGAAAGTGAATTACTATCTCTCTAATCTGAAAGGCACTCTGGGGCTGTGAATCTATT	1030
QY	844	CTTGAATATATGACAGAAAGAGGAGCATATGAGACCAATAGCCGAAGATTTCCCTTGAA	903
Db	1031	CTGGAGATTAATGACATTAAGAGAGCTCAAAATATCAAGAAATTCACAGAAAGATCCCTTGAG	1090
QY	904	ATCTTGGCACACATGGCTGTGGTTGGAAAGCTGATTGGAAAAAGAGGACAGAAATTGAG	963
Db	1091	ATTTTAGCTCATTAACTCTTGTGGAGCTGTATTGTGTAAAGAAAGAAATCTTAAA	1150
QY	964	AAAAATTGAACATGAACAGGAGCAAGATTAACATCTATCTTTGCAGATTTGAGATA	1023
Db	1151	AAAAATTGACAGACAGACAGACACTAAATATCAGATATCTCCATTTGACAGAAATGACCTG	1210
QY	1024	TACACCCGAGAAAGAACATCATCTGTGAAGGGCACAGTTGAGGGCTGTGCCAGTGCAG	1083
Db	1211	TATTAATCCAGAAACGCACTATTACAGATTAAAGCAATGTGGACATGTGGCCAAAGCTGAG	1270
QY	1084	ATGAGATTATATGAAGAGCTGCGTAGAGGCTTTGAAATATGATATGCTGCTTTAACC	1143
Db	1271	GAGGAGATCTAGAGAAATATCAGGAGTCTTATGAAATGATATATGCTTATGATCTT	1330
QY	1144	CACTCCGAT-----ACTTCCCA	1162
Db	1331	CAGGCAATTTAATCTCGATTAATATGAAAGCCTTGGTCTGTGCCACCAACCTTCA	1390
QY	1163	GCCTGTACCCCATACACAGTTTGGCCGCTTCCGATCATCACTC-----	1208
Db	1391	GGGATGCACTCTCCACCTCAAGGGGCCCTTCAAGCCATGACTCTCTCCCAACCGCAGTTT	1450
QY	1209	--TTATCCAGAGAGAGATTGTGATCTCTTATCTCCACCAACCAAGCTGTGTGGCCCATC	1266
Db	1451	GAGCATCAAGAAACGAGACTGTTCATCAGTTTATCCAGCTCTATACAGTGGGTGCATC	1510
QY	1267	ATCGGAGAGAGGGGCAACATCAAAACGCTGGCGAATTTGCCCGAGCTTATCAAG	1326
Db	1511	ATCGGAGAGAGGGCCAGACATCAAGACAGCTTTTCGCTTGTGGAGCTTCAATTGAG	1570
QY	1327	ATTGCCCCGCGAAGGCCACAGCTCAGCGAAAGATGATCATATCAACGGGGCACCG	1386

Db	1571	ATTGCTCCAGCGGAGAGACCCAGATGTCTAAAGTGAGAGATGGTGTATATACCTGACACACCA	1630
Qy	1387	GAGGCCAGTTTCAGAGGCCAGGAGACGATCTTTGGGAAACTGAAAGAGAGAAAATTCTT	1446
Db	1631	GAGGCTCAGTTCAAGGCTCAGGGAGAGAAATTTATGGAAAAATTAAGAAAGAAAATTCTGT	1690
Qy	1447	AACCCCAAGAGAAAGTGAAGCTGGAAAGCCATATCAGATGTGCCCTTCCACAGCTGGC	1506
Db	1691	AGTCTCAAAAGAGAGGTGAAGACTTGAAGCTCATATCAGATGTGCCATCTTTGCTGTGGC	1750
Qy	1507	CGGGTATTGGCAAAAGGTGGCAAGACCGGTAAAGAACTGACAGAACTTAACCACTGCAAA	1566
Db	1751	AGATTATATGGAAAAAGAGAGCCAAACGGTGAATGAATTTGAGAAATTTGTCAAGTGCAGAA	1810
Qy	1567	GTCATCTGTCGTCGCTGACCAAAACCGCAGATGAATAAGAGAGATGATCTGTCAGAATATC	1626
Db	1811	GTGTGTTTCCTCGTGACCAAGACACCTGATGAGATATCAAGAGGTGTGTCCAAATTA	1870
Qy	1627	GGGCACTTCTTTGTATGCCAGACTGCAACGGCAAGATCAGGAAATTTGTCAACAGGTG	1686
Db	1871	GGTCACTTCTATGTCTTCCAGGTTGCCAGAGAAAATTCAGGAAATTTGTACTCAGGTA	1930
Qy	1687	AAGCAGCAGAGCAGAAATACCTTCAGGAGATGCGCTCAC	1726
Db	1931	AAGCAGCAGCAGCAACAGAGGCTCTGTGAAGTGGACAC	1970

RESULT 9

US-09-899-651-4
; Sequence 4, Application US/09899651

Patent No. 6576756

; GENERAL INFORMATION:

APPLICANT: Chen, Rao-Iseng

APPLICANT: Tsang Solomon

APPLICANT: Stockert, Elis

APPLICANT: Jager, Elke

APPLICANT: Knuth, Alexander

APPLICANT: Old, Lloyd J.

TITLE OF INVENTION: Isolating

TITLE OF INVENTION:	ASSOCIATION:
;	

FILE REFERENCE: IJND 5538

CURRENT APPLICATION NUMBER

CURRENT FILING DATE: 2001

PRIOR APPLICATION NUMBER:

PRIOR FILING DATE: 1998-04

; NUMBER OF SEQ ID NOS: 8

; SEQ ID NO 4

LENGTH: 4159
TYPE: DATA

ORGANISM: Homo sapiens

FEATURE:

US-09-899-651-4

Query Match	21
Query Match	24

Best Local Similarity 64

Matches 113; Conservativ

73 ATGAACAAGCTTAA

Db 251 ATGACCAACTGTA

QY 133 CTCTTTGGGGACAG

2
0
1
2
3
4
5
6
7
8
9

DB 311 AIC11CAAGGACGC

193 TTTGGTGGACTACCC

— F
— U
— C

Db 371 TTCGTGACTGCCC

QY 253 GTGGAATTGCATGG

Query Match	21.2%;	Score 695.6;	DB 4	Length 4159;
Best Local Similarity	64.8%;	Pred: No.2,5e-184;		
Matches 1115;	Conservative	0;	Mismatches 539;	Indels 66; Gaps 3
Qy	73	ATGAAACAGCTTTACATCGGGAACTGAGCCCCGCCCGTACCGCCGACGACACTCCGGCAG	132	
Db	251	ATGAAACAACTGTATATCGGAAACCTCACGCGAAGAGCGCCGCCCTTCGGACCTAGAAAGT	310	
Qy	133	CTCTTTGGGAGACGGAAGCTGCCTCCGCGGGGACAGATCTGTGATGTCGGCTACCGC	192	
Db	311	ATCTTCAAAGACGCCCAAGATCCCGGTGTGCGGACCTTCTGATGAAGACTGCTACCGG	370	
Qy	193	TTTGTGACTACCCCGACCAAGAACTGGGGCATTCGCGCATGAGAGCCCTCTCGGGTAA	252	
Db	371	TTGTGTGACTCCCGGACGAGAGCTGGCCCTTAAGGCATCGAGGCGCTTTCAGGTAA	430	
Qy	253	GTGGAATTGCGATGGAAAAATCATGGAAGTTGATTACTCAGTCTCTTAAAAAGCTAAGAGC	312	

Db 431 ATGAACTGACGGGAAACCCATGAAAGTTGAGCACTCGGTCCCAAAAGGCAAGAGATT 490
Qy 313 AGGAAATTCAGATTGAAACATCCCTCCATCCTGCAATGGGAGAGGTTGATGACATT 372
Db 491 CGGAAATTCAGATTGAAATATCCGCTCATTTACAGTGGAGAGTGTGATGATTGA 550
Qy 373 TTGGCTCAATATGAGGACAGTGGAGATGTGAAACAAGTCAACACAGACAGAAACCGCC 432
Db 551 CTAGTCCAGTATGAGAGTGTGAGAGCTGTGAGCAAGTGAACATGACTCGAATCTGA 610
Qy 433 GTTTCACAGTCAATATGCAACAGAGAAAGCAAAATAGCCATGAGAGAGTAAAGC 492
Db 611 GTTGAATAGTAACTATTCAGTAAAGCAAGTGAACAGCACTGACAACTGAAAT 670
Qy 493 GGGCATCAGTTGAGAGTACTCTCCCTCAAGATTTCTTACATCCCGATGAAGAGTAC 552
Db 671 GGATTTCAATTGAGAAATTTACCTTGAAAGTAAAGCTTATCCCTGATGAATATGACCC 730
Qy 553 TCCCTTGGCCCCCTCAGGAGCCCAAGCGTGGGAGCACTCTTCCCGGAGCAAGGC--- 609
Db 721 CAGCAAAACCCCTTGACAGAGCCCGAGAGTGCCTGGGAGAGGGGCTCTCA 790
Qy 610 -----CAGCCCTGGGGGCACTTCTCAAGCCAGACAGATTGATTTCCCGTGGCATC 663
Db 791 AGGCAAGGGGTCTCCAGGATCCGTATCAAGCAGAAACCATGATTTGGCTCTGGCCTG 850
Qy 664 CTGGTCCCCCAGTTGTTGGTCCATCTCGAAGAGAGAGGCTTGAACCTTAAAGAC 723
Db 851 CTGGTCCCCCAGTTGTTGGAGCCATCAAGAGAAAGAGGAGCCATTTGGAGAC 910
Qy 724 ATCACTAGCAGAGCCAGTCCCGGATATCCATAGAAAGAGAACTTGGAGCTGCA 783
Db 911 ATCAACAAAGAGACCCAGCTTAAATGATGATCCATAGAAAGAGAGAGGAGCTCT 970
Qy 784 GAGAGCTGTCACTCATCCATCCCAAGAGGAGCTTGAAGATGCGGATGAT 843
Db 971 GAGAGCTGTCACTCATCTCTCACTCTGAAGGACCTCTGGGGCTGTAAGTCAAT 1030
Qy 844 CTGAATATCATGAG 903
Db 1031 CTGAAGATATGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1090
Qy 904 ATCTTGGCACAATGAGCTGTGTAAGAGATGTAAGAGAGAGAGAGAGAGAGAG 963
Db 1091 ATTTAGCTCATATTAATTTGTTGAGAGCTTATTTGTAAGAGAGAGAGAGAGAG 1150
Qy 964 AAAATTAAGATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1023
Db 1151 AAAATTAAG 1210
Qy 1024 TACAACCCGAG 1083
Db 1211 TATATCAAG 1270
Qy 1084 ATGAGATTAAG 1143
Db 1271 GAGAGAGATGAG 1330
Qy 1144 CACTCCGAGT-----ACTTCTCA 1162
Db 1331 CAGAGCATTAAATTCCTGATTAATCTGAGAGCTTGGGTCTGTTCCACCCACTTCA 1390
Qy 1163 GCTGTAACCCCATTCACCAATTTGGCCCTCCGATCACTC----- 1208
Db 1391 GGGATGCACTCCCACTCAAGGCCCCCTTCAAGCAATGCTCTCCCTACCCGAGTTT 1450
Qy 1209 --TTATCCAGAGAGAGATTTGAATCTTTCATCCCAACAGAGGTGGGAGCAGC 1266
Db 1451 GAGCATATCAGAAAG 1510
Qy 1267 ATGGGAG 1326

Db 1511 ATCGGCAAG 1570
Qy 1327 ATTGCCCTCGGAG 1386
Db 1571 ATTGCTCCAGCGAG 1630
Qy 1387 GAGGCCAGTTCAAGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1446
Db 1631 GAGGCTCAGTTCAAGGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1690
Qy 1447 AACCCAAAG 1506
Db 1691 AGCTTAAG 1750
Qy 1507 CGGATGATTTGCAAG 1566
Db 1751 AGAGTTATTTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1810
Qy 1567 GTCATGCTGCTCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1626
Db 1811 GTTGTGTCTCTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1870
Qy 1627 GGGCAGCTCTTGTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686
Db 1871 GGTCACTTCTATGCTTGCAGAGTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAG 1930
Qy 1687 AAGCAG 1726
Db 1931 AAGCAG 1970

RESULT 10
US-09-643-597-175
? Sequence 175, Application US/09643597
? Patent No. 6426072
? GENERAL INFORMATION:
? APPLICANT: Wang, Tongtong
? APPLICANT: Fan, Liqun
? APPLICANT: Kalos, Michael D.
? APPLICANT: Bangur, Chaltanya S.
? APPLICANT: Hoeken, Nancy
? APPLICANT: Fanger, Gary R.
? APPLICANT: Li, Samuel X.
? APPLICANT: Wang, Aljun
? APPLICANT: Skeiky, Yasir A.W.
? APPLICANT: Henderson, Robert A.
? APPLICANT: McNeill, Patricia D.
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
? FILE REFERENCE: 210121.455C11
? CURRENT APPLICATION NUMBER: US/09/643,597
? NUMBER OF SEQ ID NOS: 369
? SOFTWARE: PaetsSeq for Windows Version 3.0
? SEQ ID NO 175
? LENGTH: 4181
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: unsure
? LOCATION: (3347)
? OTHER INFORMATION: n=A,T,C or G
? NAME/KEY: unsure
? LOCATION: (3502)
? OTHER INFORMATION: n=A,T,C or G
? NAME/KEY: unsure
? LOCATION: (3506)
? OTHER INFORMATION: n=A,T,C or G
? NAME/KEY: unsure
? LOCATION: (3520)
? OTHER INFORMATION: n=A,T,C or G
? NAME/KEY: unsure
? LOCATION: (3538)

```

OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (3549)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (3646)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (3940)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (3968)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (3974)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (4036)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (4056)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (4062)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (4080)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (4088)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (4115)
OTHER INFORMATION: n=A,T,C or G
US-09-643-597-175

```

```

Query Match      21.2%; Score 695.6; DB 4; Length 4181;
Best Local Similarity 64.8%; Pred. No. 2.6e-184;
Matches 1115; Conservative 0; Mismatches 539; Indels 66; Gaps 3;

```

```

73 ATGACAGCTTACATCGGGAACCTGAGCCCGCCGTCACCGCGACGACTCCGCGAG 132
251 ATGACAACTGATATTCGGAACCTCAGGAGAACGCGCCCTCGGACTAGAAAGT 310
133 CTCTTTGGGAGCAGAAAGCTGCGCGGAGCAGGTCTCTGTAAGTCCGCTACGCG 192
311 ATCTTCAGGAGCGCCAAAGATCCCGGTGTCCGGAACCTTCTGTGTAAGTCTGCGG 370
193 TTGCTGACCTACCCCGACGAGCTGGGCCATCGCGCCATCGAGACCTCTCGGCTAAA 252
371 TTGCTGACTGCGCGGAGAGAGCTGGGCCCTCAAGGCCATCGAGGCGCTTTCAGGTAAA 430
253 GTGGAATTGATGGGAAATCATGGAAGTTGATTACTCACTCTTAAAGCTTAAGAGC 312
431 ATGGAATCGACGGAAGAACCATAGAGTTGAGCACTCGTCCCAAAAAGCGAAAGATT 490
313 AGGAAATTCAGATTGGAACATCTCTCTCACTGAGAGTGGAGGTGTGAGACTT 372
491 CGGAACTTCAGATACGAAATATCCGCTCATTTAAGTGGAGGTGCTGGATATTTA 550
373 TTGCTCAATATGGGACAGTGGAGATGTGGAACTGCAACAGACACAGAAACCGCC 432
551 CTAGTCCAGATGAGTGGAGTGGAGAGCTGTGAGCAAGTGAACCTGACTCGGAAACTGCA 610
433 GTTGTCAACGTCAATATGCAACAAGAGAGAGCAAAATAGCATGAGAGACTAAGC 492
611 GTTGTAAATGTAACTATTCAGATAGAGACCAAGCTAGACAAAGACATAAAGTAAAT 670
493 GGGCATAGTTTGAAGACTACTCTTCAAGATTTCTTAACATCCCGGATGAAGAGTAGC 552
671 GGATTCAGTTAGAGAAATTTCACTTGAAGATAGCTTAATATCCCTGATGAATAGCGCC 730
553 TCCCTTGGCCCTCAGCGAGCCAGAGGTGGGAGCACTCTTCCCGGAGCAAGGC-- 609

```

```

731 CAGCAAAACCCCTTGACAGAGCCCGAGGTGCGCGGGCTTGGGACAGAGGCTCTCA 790
610 -----CACGCCCTGGGGGACCTTCTCAGGCGCAGACAGATTGATTTCCGCTGGATC 663
791 AGGAGGGGCTTCCAGATCCGATCCAGAGAGAACCATGATGATTTGCTGTGGCTG 850
664 CTGATCCCGACAGTTGTTGGTCCATCATCGGAAGAGGGCTTAAACATTAAGAAC 723
851 CTGATCCCGACAGTTGTTGGTCCATCATCGGAAGAGGGCTTAAACATTAAGAAC 910
724 ATCACTAAGCAGACCTGCGGATGATATCCATAGAAAAGACCTGAGAGCTGCA 783
911 ATCACTAAGCAGACCTGCGGATGATATCCATAGAAAAGACCTGAGAGCTGCA 970
784 GAGAGGCTGACCATCATGACCCAGAGGGAGCTTGAAGACTGCGCATGAT 843
971 GAGAGGCTGACCATCATGACCCAGAGGGAGCTTGAAGACTGCGCATGAT 1030
844 CTGGAATCATGACAGAAAGGACAGATGACCAATAGCCGGAAGATTTCTGAAA 903
1031 CTGGAATCATGACAGAAAGGACAGATGACCAATAGCCGGAAGATTTCTGAAA 1090
904 ATCTTGGACACATGCTGTTGGTGAAGACTGATTTGGAAGAGGAGGATTTGAAG 963
1091 ATTTAGCTCATTAATACCTTGTGAGCTTTATTTGGTAAGAGAGAAATCTTAA 1150
964 AAATTTGAACATGAACAGGAGCAAGATTAACATCTCATCTTGAAGATTTGAAG 1023
1151 AAATTTGAACATGAACAGGAGCAAGATTAACATCTCATCTTGAAGATTTGAAG 1210
1024 TACAAACCGGAAAGAACCATCATCTGTGAAGGACAGTTGAGGCTGTGCTGAG 1083
1211 TATTAATCCAGAACGCACTATTAAGATTAAAGGCAATGTGAACAGTGCACAAAGTGA 1270
1084 ATAGAGATTATGAAGAGCTGGTGAAGGCTTTGAAGATGATGCTGCTGAAC 1143
1271 GAGAGATCATGAAGAAATCAAGAGGCTTTATGAAGATGATGCTGCTGAAC 1330
1144 CACTCGGAT-----ACTTCTCA 1162
1331 CAGGACATTTAATCTGATTAATCTGAAGCCTTGGGTGTCTCCACCACTTCA 1390
1163 GCCTGACCCCATCAACAGTTTGGCCGTTCCGATCATCATC----- 1208
1391 GGGATCCACCTCCACTCAAGGCCCCCTTCAAGCATGCTCTTCAACCGCAGTTT 1450
1209 --TTATCCAGAGAGATGTAATCTTCAATCCCAACCGAGCTGTGGCGCATC 1266
1451 GAGCAATCGAAGAGAGAGCTGTCATCACTTATCCAGCTCTATCACTGCTGATC 1510
1267 ATCGGAGAGAGGAGGACATCAACAGCTGGAGATTGCGCGAGCTCTATCAAG 1326
1511 ATGCGAGAGAGGAGGACATCAACAGCTTCTGCGCTTGAAGCTTCAATTAAG 1570
1327 ATGCGCGAGAGGAGGAGGAGCTGTCATCACTTATCCAGCTCTATCACTGAG 1386
1571 ATGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1630
1387 GAGGCGAGTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446
1631 GAGGCGAGTTCAGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1690
1447 AACCCAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1506
1691 AGTCTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1750
1507 CGGCTGATGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1566
1751 AGAGTTATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1810
1567 GTCACTGCTGCTGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1626

```


Db 1811 GTTGTGTCCTCTGTAACAGACCTGATGAGATGACCAAGTGTGTCATAAATACT 1870
 QY 1627 GGGCACTTTTGTCTAGACAGACTGCAAGCCGCAATGAGAAATTGACAGAGTG 1686
 Db 1871 GGTCACCTTCTATGCTTGTCCAGGTGTCCAGAGAAAATTGAGAAATTCTGACTAGGTA 1930
 QY 1687 AAGCAGCAGAGAGAGAAATACCTGAGGAGTGCTCTGAC 1726
 Db 1931 AAGCAGCAGCAGCAGAGAGAGGCTGTGCAAGTGAGCCAGC 1970

RESULT 11

US-09-480-884A-175
 / Sequence 175, Application US/09480884A
 / Patent No. 6482597
 / GENERAL INFORMATION:
 / APPLICANT: Wang, Tongtong
 / APPLICANT: Fan, Liqun
 / APPLICANT: Hosken, Nancy A.
 / APPLICANT: Kalos, Michael D.
 / APPLICANT: Fanger, Gary R.
 / TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 / FILE REFERENCE: 210121.455C6
 / CURRENT APPLICATION NUMBER: US/09/480, 884A
 / CURRENT FILING DATE: 2001-08-27
 / NUMBER OF SEQ ID NOS: 330
 / SOFTWARE: FastSeq for Windows Version 3.0
 / SEQ ID NO 175
 / LENGTH: 4181
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (1)_(4181)
 / OTHER INFORMATION: n=A,T,C or G
 US-09-480-884A-175

Query Match 21.2%; Score 695.6; DB 4; Length 4181;
 Best Local Similarity 64.8%; Pred. No. 2.6e-184;
 Matches 1115; Conservative 0; Mismatches 539; Indels 66; Gaps 3;

QY 73 ATGAACAAGCTTTTATCATCGGGAACCTGAGCCCGCTCAACCGGAGACCTCCGGAG 132
 Db 251 ATGAACAAGCTGTATATCGGAACCTCAGGAGAAACCGCCCTCGGACTTGAAAGT 310
 QY 133 CTCTTTGGGAGCAGAAAGCTGCCCCCTGGCGGAGAGGTCTGCTGAAGTCGGCTACGCC 192
 Db 311 ATCTTCAAGAGACGCGCAAGATCCCGGTGTGCGGACCTTCTGCTGGAAGACTGGCTACGCG 370
 QY 193 TTGCTGACTACCCGACGAGAACTGGGCCATCGCGGCATCGAGACCTCTCGGGTAA 252
 Db 371 TTGCTGACTGCCCGGACGAGAGTGGGCCCTCAAGGCCATCGAGGGCTTTTCAGGTAA 430
 QY 253 GTGGAATTCGATGGGAAATCATGGAATGATTACTCAGTCTTAAAGCTAAGAGC 312
 Db 431 ATGAAGCTGACGCGGAACCCATGAGATTGAGCACTGGTCCCAAAAAGCAAGAT 490
 QY 313 AGGAAATTCGATTTGAAACATCCCTCTCACTGAGTGGGAGGTGTTGATGACTT 372
 Db 491 CGGAACCTTCAGTACGAAATATCCGCCCTCATTTACAGTGGGAGGTGTGATGATT 550
 QY 373 TTGCTCAATATGGGACAGTGGGAATGTGGAACAAGTCAACACAGACAGAAACCGCC 432
 Db 551 CTAGTCCAGTATGAGAGTGTGGAGGTGTGACCAAGTGAACCTGACTGGAAACTGCA 610
 QY 433 GTTGTCAAGCTCATATATGCAACAGAGAGAGCAAAATAGCCATGAGAGCTAAGC 492
 Db 611 GTTGTAAATGTAACCTATTCCAGTAAAGCAAGAGCTAGCAACACTAGCAACTGAA 670
 QY 493 GGGCATCAGTTTGAAGACTACTCTTCAAGATTTCTCACTCCCGAGTGAAGGTGAGC 552
 Db 671 GGATTTCAAGTAAAGATTTCACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 730

QY 553 TCCCTTCCGCCCCCTGAGGAGCCAGCGTGGGAGCAACTCTTCCGGAGAGAGC--- 609
 Db 731 CAGCAAAACCCCTTCCAGAGCCCGAGGTGCGCCGGGGCTTGGGAGAGGGGCTCTCA 790
 QY 610 -----CAGGCCCTGGGGGACCTTCTCAGGCAAGACATTTTCCGCTGGGATC 663
 Db 791 AAGCAGGGGTCTCCAGGATCCGTATCCAGGAGAAACATGATTTGCTCTGCGCCTG 850
 QY 664 CTGTCCTCCACCCAGTTTGTGTGCATCATCGGAAAGAGGGCTTGACCTAAGAAC 723
 Db 851 CTGTTCCACCCATTTGTTGAGACCATATGAGAAAGAGTCCACATTCGGAAC 910
 QY 724 ATCTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTCTGAGCTGCA 783
 Db 911 ATCAACCAACAGACCAGCTTAAATTCAGTCCACCGTAAGAAATAGCGGGGGTGTCT 970
 QY 784 GAGAGCCGTGACCATCATGCGCCACCCAGAGGGGACTTGAAGCATGCCAGATTT 843
 Db 971 GAGAGTGAATTAATCTTCTCTAATCTGTAAGGCACTCTGCGCTTGTAAATCTTAA 1030
 QY 844 CTGGAATCATGAGAAAGAGGAGATGAGACCAACTAGCCGAGAGATTCCTGAAA 903
 Db 1031 CTGAGATTAATGCAATAGGAGCTGAGATTAATAATCAAGAGAGATCCCTTGAAG 1090
 QY 904 ATCTTGGCAACAATGCTTGGTGGAGAGCTGATTTGAAAAGAGGAGAAATTTGAAG 963
 Db 1091 ATTTAGCTCATTAATTAATCTTGTGAGAGTCTTATTTGTAAGAGAGAAATCTTAA 1150
 QY 964 AAAATTAACATGAAACAGGAGCAAGTAATCAATCTTGTGAGAGATTTAGCATTA 1023
 Db 1151 AAAATTAAGCAAGCAAGACACTTAAATCAAGATTCATGAGAGATTTGAGCTG 1210
 QY 1024 TACAACCCGGAAGAAACCATCATCTGTGAAGGCAAGTTGAGGCTGCGGAGCTGAG 1083
 Db 1211 TATATCCAGAGCCACTATTAAGTTTAAAGCAATGTTGAGATGTCGCAAGCTGAG 1270
 QY 1084 ATGAGATTAATGAAGAGCTGCGAGGCTTTGAAATGATGCTGCTGTTAACACC 1143
 Db 1271 GAGAGATCAATGAAGAAATCAAGGAGCTTATGAATGATTTGCTCTATGATCTT 1330
 QY 1144 CACTCGGAT-----ACTTTCGA 1162
 Db 1331 CAGGACATTAATTTCTGATTAATATGAAACGCTTGGGTCTGTTCCACCACTTCA 1390
 QY 1163 GCCTGACCCCATCAAGATTTGGCCGTTCCCGCATCACTC----- 1208
 Db 1391 GGGATGCCACTTCCCACTCAGGGCCCTTCAAGCATGACTCTCTCCATACCGCAATTT 1450
 QY 1209 --TTATCAGAGAGAGATTTGATCTCTTATCCCAACCGAGCTGTGGGCGCATC 1266
 Db 1451 GAGCAATCAAGAAACGAGAGCTGTTCACTGATTAATCCAGCTGATCAAGTGGGCATC 1510
 QY 1267 ATGCGGAAGAGGGGGGACATCAACACAGCTGCGGAGATTCGCGGAGCTTATCAAG 1326
 Db 1511 ATGCGGAAGAGGCGAGCAATCAACAGCACTTCTCGCTTGTGAGGCTTCAATTAAG 1570
 QY 1327 ATTGCCCTGCGGAGGCGCAGAGCTGAGGAAAGATGCTCATCAACCGGCGCACCG 1386
 Db 1571 ATTGCTCCAGCGGAGACCAAGATGCTTAAAGTGAAGTGTGATTAATCACTGAGCAACA 1630
 QY 1387 GAAAGCCAGTTCAAGGCCAGGAGCGAATCTTTGGGAACTGAAAGAGGAAACTTTT 1446
 Db 1631 GAGGCTCAGTTCAAGGCTCAGGAGAAATTTATGAAAAATTAAGAAAGAAACTTTGTT 1690
 QY 1447 AACCCCAAGAAAGAAAGTGAAGCTGGAAGGCAATATCAAGTGCCTCTTCAAGCTGCG 1506
 Db 1691 AGTCTTAAGAAAGAGTGAAGCTTGAAGCTCAATATCAAGTGCATCTTTGCTGCTGCG 1750
 QY 1507 CGGCTGATTTGGCAAGAGTGAAGACCGTGAAGCACTGAGAACTTAACAGATGAGAGAA 1566
 Db 1751 AGAGTTATTTGAAAAAGAGGCAAGCAAGCGTGAATGAATCTCAAGATTTGCAAGTGAAG 1810

QY 904 ATCTGGACAAATGGCTTGGTTGGAGACTGATTGGAAAAAGAGGAGAAATTTGAAG 963
Db 1091 ATTTAGCTCATTAATACTTTTGTGACGCTTTATGTGAAAGAAAGAAATCTTAAA 1150
QY 964 AAAATTTGAACATGAACAGGGGACCAAGATTAACATCTACTTTTGAGAGATTGAGCAT 1023
Db 1151 AAAATTTGACAAAGACAGACACTAAATACAGATATCTCCATTCAGAGAAATTTGACGCTG 1210
QY 1024 TACAAACCGGAAAGAACCATCATCTGTGAAGGCGACAGTTGAGCCGTGGCCAGTCTGAG 1083
Db 1211 TATAATCCAGAAACGACATTTACAGTTAAAGGCAATGTTGAGACATGTGCCAAAGCTGAG 1270
QY 1084 ATAGAGATTATGAGAGAGCTGCTGAGGCTTTGAAAAATGATATGCTGTTAAACACC 1143
Db 1271 GAGGAGATCATGAGAAATATCAGGAGTCTTATGAAATGATATGCTTATGAAATCTT 1330
QY 1144 CACTCCGGAT-----ACTTCTCCA 1162
Db 1331 CAAGCACTTTAATCTCTGATTAATCTGAACGCTTGGTCTGTCTCCACCACTTCA 1390
QY 1163 GCTGTATACCCCATCACCAAGTTTGGCCCGCTCCCGCATCATCATC----- 1208
Db 1391 GGGATGCCACTCCCACTCTCAGGCGCCCTTACAGCATGACTCTCTCCCTACCCGCACTT 1450
QY 1209 --TTATCCAGACAGAGAGATTGTAATCTTCAATCCCAACCCAGGCTGTGGCGCATC 1266
Db 1451 GAGCAATCAGAAACGAGAGACTGTTCAATCAGTTTATCCAGCTTATCAGTCCGCGCATC 1510
QY 1267 ATCGGAAAGAAAGGGGCGACATATTAACAGCTGCGAGATTGCGCGAGCTTATATAG 1326
Db 1511 ATCGGCAAGCAGGGGCGACATCATCAGCATCTTCTGCTTGTGGAGCTTCAATTAAG 1570
QY 1327 ATTGCCCCGCGGAAAGGCGGAGCGTCAAGCGAAAGAGATGTCATCATCACCGGCGCACCG 1386
Db 1571 ATTGCTCCAGGAGAGACCAAGATCTTAAGTGAAGATGATTAATCATCTGACACCA 1630
QY 1387 GAAGCCCAATTCAAGGCGGAGCGGATCTTTGGGAAATGAAAGAGAAATCTTCTT 1446
Db 1631 GAGGCTCAGTTCAAGGCTCAGGAGAAATTTATGAAATTAAGAAAGAAATCTTGT 1690
QY 1447 AACCCCAAGAAAGAGAGAGTGGAGCGCATATCAAGTCCCTTCCACAGCTGCG 1506
Db 1691 AGTCTTAAAGAAAGGAGGAACTTGAAGCTCATATCAGATGCTTGTGCTGCTGCG 1750
QY 1507 CGGGTGAATGGCAAGGTGGCAAGACCGTGAACGAACTGAGAACTTAACAGTGCAGAA 1566
Db 1751 AGAGTTATTGAAAAAGAGGCAAAACGTTGATGAATCTTCAAGATTTGTCAAGTGCAGAA 1810
QY 1567 GTCATCGTGCCTCGTGAACCAACGCGCATGAATGAGAGATGATCTCAGAAATTAATC 1626
Db 1811 GTTGTGTCCCTCGTGAACGACACCTGATGAGATGACCAAGTGTGTCAAAATTAATC 1870
QY 1627 GGGCACTTCTTCTGAGCCGAGATCTGACAGCGCAAGATCAAGGAAATTTGTAACAAGGTG 1686
Db 1871 GGTCACTTCTATGCTTGGCCAGGTTGCCAGAGAAAAATTCAGAAATTTGACTCAGGTA 1930
QY 1687 AAGCAGCAGAGAGAGAAATTAATCTCAGGAGGTGCGCTCAC 1226
Db 1931 AAGCAGCAGCAGAGAGAGGCTTGCAGAAATGAGCAC 1970

RESULT 13
US-09-606-421B-175
; Sequence 175, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Fan, Aijun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.

APPLICANT: Wang, Aijun
TITLE OF INVENTION: Skeiky, Jasir A.W.
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C9
CURRENT APPLICATION NUMBER: US/09/606,421B
NUMBER OF SEQ ID NOS: 358
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 175
LENGTH: 4181
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (3347)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3502)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3506)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3520)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3538)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3549)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3646)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3940)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3968)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (3974)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4036)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4056)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4080)
OTHER INFORMATION: n=A,T,C or G
FEATURE:
NAME/KEY: unsure
LOCATION: (4088)
OTHER INFORMATION: n=A,T,C or G

```

; FEATURE:
; NAME/KEY: unsure
; LOCATION: (4115)
; OTHER INFORMATION: n=A,T,C or G
US-09-606-421B-175

```

Query Match	21.2%;	Score 695.6;	DB 4;	Length 4181;
Best Local Similarity	64.8%;	Pred. No. 2.6e-184;		
Matches 115;	Conservative	0;	Mismatches 539;	Indels 66;
				Gaps 3;

QY	73	ATGAAACAAGCTTTAACAATCGGGAAACCTGAGCCCGCCGCTCACCGCCGACATCCTCCGCGAG	132
Db	251	ATGAACAAACTGTAATATCGGAAACCTCAGAGAAAGCGCCGCCCTCCGACCTAGAAAGT	310
QY	133	CTCTTTGGGGACAGAAAGTCCCTCGCGGGGACAGAGTCTCTGTAAATTCGCGCTAACGCC	192
Db	311	ATCTTCAAGGACGCAAGATCCCGGAGTCCGGGACCCCTCTCGTGTAAAGCTGGCTAACCGG	370
QY	193	TTTCGTGACATACCCCGGACCGAAACTGGGGCCATCCGGGCCATCGAGACCCCTCCGGGTAA	252
Db	371	TTTCGTGACCTGCCGACGAGAGCTGGGCCCTTAAGGCCATCGAGGCGCTTTCAGGTAAA	430
QY	253	GTGGAATTTGACATGGGAAATATCATGSAAGTTGATTTACTCACTCTCTAAAAAGCTTAAGAC	312
Db	431	ATTAAGACTGACGGGAAACCCATTAAGAGTTGAAGACTCGAGTCCCAAAAAGGCAAAAGATT	490
QY	313	AGGAAATTTGAGTTGAAACATCCCTCTCTCACTGACGTGGGAGGTGTGTGATGACCTT	372
Db	491	CGGAATCTTCAGATACGAAATATCTCCCGCTCATTTACGTGGGAGGTGCTGGATTA	550
QY	373	TTGGCTCAATATGGGACAGTGGAAATAGTGGAAACAAGTCAACACAGACAGAAACCGCC	432
Db	551	CTAGTCCAGTATGAGAGTGGTGAAGCTGTGAGCAATGAACTGACTGTCTGGAAACTGCA	610
QY	433	GTTGTCAACGTACATATNGCAACAAGAGAAAGCAAAAATAGCATGAGAAAGCTAAGC	492
Db	611	GTTGTAATGTAACTTATTCAGTAAGAACCAACTGTAAGACAAGCACTAAGCAAACTGAAT	670
QY	493	GGGCATCAGTTTGAAGAACTACTCTTCACAGATTTCTTCACTCCCGATGAAGAGTGAAC	552
Db	671	GGATTTCAGTTAAGAAATTTCACTTGAAGTAAGTCACTTATCTCTGATGAAGAGCGGCC	730
QY	553	TCCCTTGGCCCCCTCAGCCGAGCCCAAGTGGGGAGCAACTCTTCCGGGAGCAAGGC---	609
Db	731	CAGCAAAACCCCTTGCACAGACCCCGAGAGTGGCGGGGCTTGGGCAAGAGGGCTCTCA	790
QY	610	-----CAGGCCCTGGGGGCACTTCTCAGGCCACAGATTGATTTCCGCTGCGGATC	663
Db	791	AGCAGGGGCTCTCCAGATCCGTATCCAGAGAAACCATGTATTTGGCTCTGCGGCTG	850
QY	664	CTGTCTCCCAACCCAGCTTTGTGTGGCATCATCGGAAAGAGGGCTTGACCAATAAGAAC	723
Db	851	CTGGTTCCCAACCAATTTGTGGACCATCATAGGAAAGAGGTGGCACCATTCGGAAC	910
QY	724	ATCACTAAGCAGACCCAGTCCCGGGTAATATCCATATAAAAGAGAACTCTCGAGCTGCA	783
Db	911	ATCAACCAACAGACCCAGCTTAAATTCATGTCCACCGTAAAGAAATATGCGGGGCTGCT	970
QY	784	GGAAGACCTGTACCATCATCCATGCCACCCAGAGGGGACCTTCTGAAGCATGCGGATATT	843
Db	971	GGAATGTCAATTAATCTCTCTCTACTCTGAAGGCACTCTGCGGCTGTAAATCTAATT	1030
QY	844	CTTGAATATCATCAGAAAGGCGAGATGACCAAACTAGCCGAAGAGATTTCTCTGAAA	903
Db	1031	CTGGAGATTATGCAATAAGAAAGCTCAAGATATAAAATTCACAGAAAGATCCCTCTGAAG	1090
QY	904	ATCTGGCAACAATGGCTTGGTGGAGACTGATTTGGAAAAAGAGGAGAAATTTGAGG	963
Db	1091	ATTTTAGCTCATATAATTAATTTGTGGACCTCTTATTTGAAAGAAAGAAATCTTTAAA	1150
QY	964	AAAAATTGAACATGAACAGGGGACAAAGATTAACAATCTCATTTTSCAGATTTTGGACATA	1023
Db	1151	AAATTTGACACAGACACACTAAATTCAGATATCTTCATTCAGAGAAATGACGCTG	1210

```

OY      1024  TACACCCCGAAGAAGCAATCACTGTGAAGGGGACAGTTGAGGCGCTGTGCAGTGTGAG      1083
Db      1211  TATATATCCAGAAAGCACTATTATTCAGTTTAAAGGCATACTTGACATGTGCAGAAAGCTGAG      1270
OY      1084  ATAGAGATTATGAAGAAGCTGCCTGAGGCTTTGAAATGATATATGCTGGCTGTAAACCC      1143
Db      1271  GAGAGAGATCATGAAGAAATATCAGGAGAGCTTATGAAATATGATATGCTTTATGAATCTTT      1330
OY      1144  CACTCCGGAT-----ACTTCCCA      1162
Db      1331  CAAGCACAATTAAATCTCTGATTTAATCTGAAGCCTTGGGTCTGTTCACCAACCACTTCA      1390
OY      1163  GCGTGTACCCCAATCACAGATTTGGCCCGTTCCCGATCATCACTC-----      1208
Db      1391  GGGATGCTACCTCCCACTCAGAGGCCCCCTTCAGACCATGACTCCTCCCTAACCCGCACTTT      1450
OY      1209  --TTATCCAGACAGAGGATTTGTCAATCTCTTATATCCCAACCAAGAGGTGGGGGCGCATC      1266
Db      1451  GAGCAATTCAGAAACGGAGACTGTTCATCAGTTTATCCAGCTCTATATAGTCGGTGCCATC      1510
OY      1267  ATCGGAAGAAGAGGGGACACATCAATCAACAGCTGGCGAGATTCCGCGAGGCTCTATCAAG      1326
Db      1511  ATCGGCAAGAGGGCCACGACATCAACAGAGTTTCTCGCTTGTGCGAGCTTCAATTAAAG      1570
OY      1327  ATTGCCCCCTGCGAAGGCCACAGCTCAGCGAAGATGTGCATCATCAACGGGCGCACCG      1386
Db      1571  ATTGCTCCAGCGGAAAGCACCGATGCTAAAGTGAAGATGTGATTTATCACTGACACCA      1630
OY      1387  GAACCCCAAGTTCAGAGGCCACAGGACCGGATCTTTGGGAAATCGAAAGAGGAAATCTTT      1446
Db      1631  GAGGCTCAGTTCAGAGGCTCAGGGAAGATTTATGAAAAATTTAAAGAAAGAAATCTTGT      1690
OY      1447  AACCCTAAAGAAAGTAGAAGCTGAAAGCGCATATCAGAGTGCCCTCTTCCACAGCTGGC      1506
Db      1691  AGCTCTAAAGAAAGGTGTAACTTGAAGCTCATATCAGAGGCCATCTCTTGTGCTGGCG      1750
OY      1507  CGGTTGATTGGCAAAAGTGTGCACAAACCGGTGAACGAACTGCAAGACTTAACAGTGCAGAA      1566
Db      1751  AGAGTATTATGGAAGAAAGGAGGCAAAACGGTGAATGAACTTCAGAAATTTGTCAAGTGCAGAA      1810
OY      1567  GTCATCTGTCCTCGTGAACCAACGCCAATGAAATTAAGAAAGTGAATGTGTAAGATTATC      1622
Db      1811  GTTGTGTTCCTCTCGTGAACCAACCTGATGATGATGACCAAGTGTGTGTCAAAATTAATC      1870
OY      1627  GGGCACTTCTTGTCTAGCCACAGTGCACAGCGCAAGATCAGGGAATTTGTACAAAGGTTG      1686
Db      1871  GGTCACTTCTATGTCTTGGCCAGGTTGCCACGAAATAATTCAGAAATTTCTGACTCAGGTA      1930
OY      1687  AAGCAGCAGAGACAGAAATATACCTCAGGAGAGTGGCTCAC      1726
Db      1931  AAGCAGCAGCAGCAGCAGAAAGGCTTGCAAAGTGCAGCAC      1970

RESULT 14
US-09-261-855-1
/ Sequence 1. Application US/09261855A
/ Patent No. 6255055
/ GENERAL INFORMATION:
/ APPLICANT: Ross, Jeffrey
/ TITLE OF INVENTION: (THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
/ TITLE OF INVENTION: TRD-BP) AND ITS NUCLEIC ACID SEQUENCE
/ FILE REFERENCE: 960296.95131
/ CURRENT APPLICATION NUMBER: US/09/261, 855A
/ CURRENT FILING DATE: 1999-03-03
/ NUMBER OF SEQ ID NOS: 46
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 2224
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-261-855-1

```

Query Match 20.9%; Score 685; DB 3; Length 2224;
 Best Local Similarity 64.1%; Pred. No. 1.5e-181;
 Matches 1117; Conservative 0; Mismatches 560; Indels 66; Gaps 3;

QY 73 ATGACAAGCTTTAATCGGGAACTGAGCCCGCCGCTGACCGCGAGCACTCCGCG 132
 DB 131 ATGACAAGCTTTAATCGGGAACTGAGCCCGCCGCTGACCGCGAGCACTCCGCG 190
 QY 133 CTCTTTGGGGACAGGAGCTGCCCCCTGGCGGGACAGGTCCTGCTGAGCTCGGCTAA 192
 DB 191 GTATTCCGGGAGCAAAAGATCTCCACAGCGCCGCTGTTGGTAAATCCGGCTACGCC 250
 QY 193 TTCTGGACCTACCCGACCAAGCTGGGCCATCCCGCCATCGAGCCCTCTCGGGTAA 252
 DB 251 TTCTGGACCTACCCGACCAAGCTGGGCCATCGAGCCCTCTCGGGTAA 310
 QY 253 GTGGAAATTCGATGGGAAAATATGAGAAATTTGATTTCTAGTCTCTAAAAGCTAAGGC 312
 DB 311 GTAGAACTGCAAGGAAAAGCTCTAGAGATTGAACTCAGTCCCAAAAAAATAAGGAGT 370
 QY 313 AGGAAATTCAGATTTCGAATTCGATCCCTCTCAGCTGCGAGGAGTGTGGATGAGCTT 372
 DB 371 CGGAAATTCAGATTTCGATCCCTCTCAGCTGCGAGGAGTGTGGATGAGCTT 430
 QY 373 TTGGCTCAATATGAGCAAGTGGAGATGTGGAACAAGTCAACAGACACAGAAACGCGC 432
 DB 431 CTGGCTCAATATGAGCAAGTGGAGATGTGGAACAAGTCAACAGACACAGAAACGCGC 490
 QY 433 GTTGTCAACCTCAATATGAGCAACAGAGAAAGCAAAAATAGCTATGAGAGCTAAGC 492
 DB 491 GTTGTCAACCTCAATATGAGCAACAGAGAAAGCAAAAATAGCTATGAGAGCTAAGT 550
 QY 493 GGGCATGATTGGAATCTACTCTTCAAGATTCTTCAATCCCGGATGGAAGAGTGAAC 552
 DB 551 GGGCATGATTGGAATCTACTCTTCAAGATTCTTCAATCCCGGATGGAAGAGTGAAC 610
 QY 553 TCCCTCTC-----GCCCTCTCAGCGAGCCGAGCGTGGGAGCACTCTCCGGGAG 603
 DB 611 CAAGTCTCTGAGATGGGCGTCTGAGAGCTTGGGCTGGGGCCAGCCCGGCAAGG 670
 QY 604 CAAGGCGAGCCCTGGGGGCACTTCTCAGGCCAGACAGATTGTTCCGCTGGGATC 663
 DB 671 TCGCCCGTGGCAGAGGGGCTCCAGCCAGAGAGCAGTGGGACATCCCTCTCCGCTC 720
 QY 664 CTGGTCCCGCAGCTGTTGTTGGTCCATCATCGGAAAGAGGGCTTGAACATTAAGAC 723
 DB 731 CTGGTCCCGCAGCTGTTGTTGGTCCATCATCGGAAAGAGGGCTTGAACATTAAGAC 790
 QY 724 ATCACTAAGCAGACCCAGTCCCGGTAATATCCATGAAAAGAACTTGGAGCTGCA 783
 DB 791 ATCACTAAGCAGACCCAGTCCCGGTAATATCCATGAAAAGAACTTGGAGCTGCA 850
 QY 784 GAGAGCTCTGTCACATCCATGTCACCCAGAGGGGACTTCTGAAGATCCGCTGATTT 843
 DB 851 GAGAGCTCTGTCACATCCATGTCACCCAGAGGGGACTTCTGAAGATCCGCTGATTT 910
 QY 844 CTGGAATCATGAGAAAAGAGCAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAA 903
 DB 911 TTGGAATCATGAGAAAAGAGCAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAA 970
 QY 904 ATCTTGGCACAACATGCTTGGTGGAAAGCTGATTGAAAAGAGCAGAAATTTGAAG 963
 DB 971 ATCTTGGCACAACATGCTTGGTGGAAAGCTGATTGAAAAGAGCAGAAATTTGAAG 1030
 QY 964 AAAATTTGAACATGAAAAGAGCAGATGAGACCAAACTAGCTTCTTGAAGATTTGACATA 1023
 DB 1031 AAGGTGAGCAGAGACCAAGAGCAGAAATCACCATCTCATGCTCAGAGACCTCAGGCTC 1090
 QY 1024 TACAACCCCGGAAAGAACATCTGTTGAGAGGGGACAGTTGAGGCCCTGTGCTGCTGAG 1083
 DB 1091 TATTAACCTGAGAGGACCATCTGTTGAGAGGGGCGCCATTGAGAACTGTTGACAGGGCGAG 1150
 QY 1084 ATTAGAATTATGAAGAGCTGGCTGAGGCCCTTTGAAAATGATATGCTGGCTGTTAACACC 1143

DB 1151 CAGGAGATCATGAGAAAAGTTCTGAGAGGCTTACGAGAACGAGCTGGCCGATGAGCTTG 1210
 QY 1144 CACTCCGAGTACTTCTCCAGCCTGTACC-----CCC 1174
 DB 1211 CAGTCCAGCTCATCTCCGAGCTTAACTGGCTGTGTAGTCTTCTCCAGCTTCATCC 1270
 QY 1175 ATCACAAGTTGGCCCTGTTCCCGATCATCT----- 1207
 DB 1271 AGCGCTGCTCCCTCTCCAGAGTGTCACTGGAGGCTGCTCCATATGCTCTTCAAG 1330
 QY 1208 -CTTATCCAGAGAGGAGATTTGATCTCTTCAATCCCAACCCAGGCTGTGGGGCATC 1266
 DB 1331 CAGGCTCCGAGCAGAGAGATGTACAAAGTTTATCCCGCCAGCAGCTGTGGGCCCATC 1390
 QY 1267 ATCCGGAAGAGAGGGGACACATCAATAACAGCTGCGAGATTCGCGGAGCCCTTATCAAG 1326
 DB 1391 ATTGGCAAGAGAGGCGCAGACATCAATACAACTCTCCGCTTTCGCGAGCGCTTCATCAAG 1450
 QY 1327 ATTGCCCCCTGCGGAAGGCCCGAGCTCAGCGAAAGATGTCATCATCACCGGCCACCG 1386
 DB 1451 ATTGCTCCACGAGAAACCTGACTCAAAAGTTGGAATGTCATCATCACTGAGACCCCA 1510
 QY 1387 GAAGCCAGTTCAAGGCCCGAGGACGATCTTGGGAACTGAAAGAGAAAATTCTTT 1446
 DB 1511 GAGCTCAAGTTCAAGGCTCAGGAGAAAGATTTATGGCAAACTAAAAGAGAAATTTCTTT 1570
 QY 1447 AACCCCAAGAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGGCCCTCTCCAGAGCGGC 1506
 DB 1571 GGTCCCAAGAGAGAAAGTAAAGCTAGACACCACTACATGAGGTTCCGCTTTCAGAGCGGC 1630
 QY 1507 CGGCTGATTGGCAAAAGGTGGCAAGACCGTGAACCACTGAGAACTTAACCACTGACAGAA 1566
 DB 1631 CGGCTCAATCGCAAAAGCGGCAAAACCGTGAATGAGCTGCAAGAACTTGACGAGCTGAG 1690
 QY 1567 GTCATGCTGCTGTGACCAAAAGCCGATGAAATATGAGAAAGTGAATCTGAGAAATTATC 1626
 DB 1691 GTGCTATGTCGCAAGAGCCAGACCCCGATGAGAAACCAAGCATTTGTTAATATATC 1750
 QY 1627 GAGCATTCTTCTGACGACAGCTGCACAGCGCAAGATCAGGGAATTTGTAACAACAGGTG 1686
 DB 1751 GGAATTTCTATCCAGCCAGATGAGCTCAGCGGAAGATCCGAACATCTCGGCTCAAGTT 1810
 QY 1687 AAGCAGCAGAGCAAGAAATACCTCAGGAGTGCCTCAGAGCGCAGCAAGTGAAGGCTCC 1746
 DB 1811 AAGCAGCAGCAGCAAGAAAGGACAGAGCAACTCGGCCAGGCAAGAGATGACCCGCG 1870
 QY 1747 CAC 1749
 DB 1871 CCC 1873

RESULT 15
 US-09-061-709-7
 ; Sequence 7, Application US/09061709B
 ; Patent No. 6297364
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Yao-Tseng
 ; APPLICANT: Gure, Ali
 ; APPLICANT: Tseng, Solam
 ; APPLICANT: Stockert, Elisabeth
 ; APPLICANT: Jager, Elke
 ; APPLICANT: Knuth, Alexander
 ; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer Associated
 ; FILE OF INVENTION: Antigens, The Antigens Per Se, And Uses Thereof
 ; FILE REFERENCE: IUD 5538
 ; CURRENT APPLICATION NUMBER: US/09/061,709B
 ; CURRENT FILING DATE: 1998-04-17
 ; SEQ ID NO 7
 ; LENGTH: 1946
 ; TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:

Query Match 17.8%; Score 585.4; DB 3; Length 1946;
Best Local Similarity 64.1%; Pred. No. 1.2e-153;
Matches 964; Conservative 0; Mismatches 481; Indels 60; Gaps 3;

303 GCTAAGGAGAGAAATTCAGATTGCAACATCCCTCTCTCACTGCGAGTGGAGTCTT 362
104 GGTTAGAGAGCCGGAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTC 163
363 GGTAG 422
164 GAG 223
423 AGAAG 482
224 TGAAG 283
483 GAG 542
284 GAG 343
543 -----AAG 593
344 GAG 403
594 TTCCGAG 653
404 CCGAG 463
654 GCTGAG 713
464 CTTGAG 523
714 GATTAAG 773
524 CATCCGAG 583
774 TGAAG 833
584 AGGTGAG 643
834 CCGAG 893
644 TAAG 703
894 TCCGAG 953
704 TCCGAG 763
954 AATTTGAG 1013
764 GAACTGAG 823
1014 TTGAG 1073
824 CTTTACCTTTTACAG 883
1074 CAGTGTGAG 1133
884 CAGGAG 943
1134 TGTAAACCCACTCCGAGATCTTCTCAGAGCTG-----TAC 1170
944 CATAGAGCTCTCACTGATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1003
1171 CCCGATCAG 1207
1004 ATCCAG 1063
1208 -----CTTATCCAG 1262

1064 TATCAGAGCTCCGAG 1123
1263 CATATGAG 1322
1124 CATATGAG 1183
1323 CAAGATGAG 1382
1184 CAAGATGAG 1243
1383 ACCGAG 1442
1244 GCGAG 1303
1443 CTTTAACCCAAAG 1502
1304 CTTTGTCTCCAG 1363
1503 TGGCGAG 1562
1364 TGGCGAG 1423
1563 AGAGATCATCTGCTCTGAG 1622
1424 TGAAGTGTAGTACCAAG 1483
1623 TATGAG 1682
1484 CATCGAG 1543
1683 GGTGAG 1742
1544 GGTGAG 1603
1743 CTCCC 1747
1604 AGCCC 1608

Search completed: July 24, 2004, 16:48:44
Job time : 248.862 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 24, 2004, 13:04:51 ; Search time 1413.28 Seconds

(without alignments)
11355.361 Million cell updates/sec

Title: US-09-270-437d-8

Perfect score: 3283

Sequence: 1 ggcagcgagagcgagcgagga.....aaccttgaaatgttattc 3283

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3216467 seqs, 244419694 residues 6432934

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US05_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3282	100.0	3283	9	US-09-899-651-8
2	3255.4	99.2	3905	10	US-09-814-353-20241
3	3253.8	99.1	3667	15	US-10-097-340-146
4	3253.8	99.1	3667	17	US-10-648-593-48
5	3143	95.7	3412	17	US-09-899-651-6
6	3052.6	93.0	3684	9	US-09-764-864-329
7	1989.8	60.6	2250	9	US-09-764-864-332
8	1561.2	47.6	1707	16	US-10-262-445-39
9	1096.8	33.4	1186	9	US-09-764-864-749
10	697.2	21.2	1740	9	US-09-735-705-347
11	697.2	21.2	1740	9	US-09-850-716A-347
12	697.2	21.2	1740	9	US-09-897-778-347
13	697.2	21.2	1740	13	US-10-007-700-347
14	697.2	21.2	1740	15	US-10-117-982-347

15	697.2	21.2	1740	15	US-10-117-982-478	Sequence 478, App
16	697.2	21.2	1740	16	US-10-313-986-347	Sequence 347, App
17	697.2	21.2	1740	16	US-10-313-986-478	Sequence 478, App
18	697.2	21.2	1743	9	US-09-897-778-447	Sequence 447, App
19	697.2	21.2	1743	9	US-09-897-778-450	Sequence 450, App
20	697.2	21.2	1743	13	US-10-007-700-447	Sequence 447, App
21	697.2	21.2	1743	13	US-10-007-700-450	Sequence 450, App
22	697.2	21.2	1743	15	US-10-117-982-447	Sequence 447, App
23	697.2	21.2	1743	15	US-10-117-982-450	Sequence 450, App
24	697.2	21.2	1743	16	US-10-313-986-447	Sequence 447, App
25	697.2	21.2	1743	16	US-10-313-986-450	Sequence 450, App
26	697.2	21.2	1799	16	US-10-313-986-485	Sequence 485, App
27	697.2	21.2	34555	15	US-10-117-982-479	Sequence 479, App
28	697.2	21.2	34555	16	US-10-313-986-479	Sequence 483, App
29	697	21.2	1740	16	US-10-313-986-483	Sequence 483, App
30	697	21.2	4159	9	US-09-899-651-4	Sequence 4, App1
31	695.6	21.2	4181	9	US-09-735-705-175	Sequence 175, App
32	695.6	21.2	4181	9	US-09-954-456-715	Sequence 715, App
33	695.6	21.2	4181	9	US-09-850-716A-175	Sequence 175, App
34	695.6	21.2	4181	9	US-09-897-778-175	Sequence 175, App
35	695.6	21.2	4181	10	US-09-466-396A-175	Sequence 175, App
36	695.6	21.2	4181	10	US-10-007-700-175	Sequence 175, App
37	695.6	21.2	4181	13	US-10-117-982-175	Sequence 175, App
38	695.6	21.2	4181	15	US-10-117-982-175	Sequence 175, App
39	695.6	21.2	4181	16	US-10-313-986-175	Sequence 175, App
40	695.6	21.2	4181	16	US-10-116-802-145	Sequence 145, App
41	694.2	21.1	1764	9	US-09-850-716A-428	Sequence 428, App
42	694.2	21.1	1764	9	US-09-897-778-428	Sequence 428, App
43	694.2	21.1	1764	13	US-10-007-700-428	Sequence 428, App
44	694.2	21.1	1764	15	US-10-117-982-428	Sequence 428, App
45	694.2	21.1	1764	16	US-10-313-986-428	Sequence 428, App

ALIGNMENTS

RESULT 1
US-09-899-651-8
? Sequence 8, Application US/09899651
? Patent No. US2002011470A1
? GENERAL INFORMATION:
? APPLICANT: Chen, Yao-Tsang
? APPLICANT: Gure, Ali
? APPLICANT: Tsang, Solam
? APPLICANT: Stockert, Elisabeth
? APPLICANT: Jager, Elke
? APPLICANT: Knuth, Alexander
? APPLICANT: Old, Lloyd J.
? TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
? TITLE OF INVENTION: Associated
? FILE REFERENCE: LUD 5538
? CURRENT APPLICATION NUMBER: US/09/899,651
? CURRENT FILING DATE: 2001-07-06
? PRIOR APPLICATION NUMBER: US/09/061,709
? PRIOR FILING DATE: 1998-04-17
? NUMBER OF SEQ ID NOS: 8
? SEQ ID NO 8
? LENGTH: 3283
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
US-09-899-651-8

Query Match 100.0%; Score 3282; DB 9; Length 3283;
Best Local Similarity 100.0%; Pred No. 0;
Matches 3283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGCGAGGAGCGAGGAGCGCGGTAACCGGCGGAGCGCGGCTCTCGGG 60
DB 1 GGCAGCGAGGAGCGAGGAGCGCGGTAACCGGCGGAGCGCGGCTCTCGGG 60
QY 61 AAGAGCGAGTATGACAAAGCTTTACATCGGAGACCTAGGCCCGCGTACCGCGAC 120


```
Db 61 AAGAGAGGATGATGAACAGCTTTTACATCGGGAACCTGAGCCCGGCTGACCGCCGAC 120
Qy 121 GACCTCCGGGACGCTCTTTGGGGACAAGAGCTGCCCTTGCGGGACAGCTCTGCTGAAG 180
Db 121 GACCTCCGGGACGCTCTTTGGGGACAAGAGCTGCCCTTGCGGGACAGCTCTGCTGAAG 180
Qy 181 TCCGGCTAAGCCTTGTGTGACTACCCGACAGAACCTGGGCAATCCGGGCAATGAGACC 240
Db 181 TCCGGCTAAGCCTTGTGTGACTACCCGACAGAACCTGGGCAATCCGGGCAATGAGACC 240
Qy 241 CTCTCGGCTAAGTGAATGTCATGGGAAAATCATGGAAGTGTATCTAGTCTCTAAA 300
Db 241 CTCTCGGCTAAGTGAATGTCATGGGAAAATCATGGAAGTGTATCTAGTCTCTAAA 300
Qy 301 AAGCTAAGAGAGCAAGAAAATTCAGATTCGAAAATCTCCCTCTCACTTGAGTGGAGGTG 360
Db 301 AAGCTAAGAGAGCAAGAAAATTCAGATTCGAAAATCTCCCTCTCACTTGAGTGGAGGTG 360
Qy 361 TTGATGATCTTTTGGCTCAATATGGGACAGTGGGAATGTGGAAACAAGTCAACACAGAC 420
Db 361 TTGATGATCTTTTGGCTCAATATGGGACAGTGGGAATGTGGAAACAAGTCAACACAGAC 420
Qy 421 ACAGAAACCGCGCTTGTCACAGCTCAATATGCAACAGAAAGAACAAAATTAAGCATG 480
Db 421 ACAGAAACCGCGCTTGTCACAGCTCAATATGCAACAGAAAGAACAAAATTAAGCATG 480
Qy 481 GAGAAAGCTAAGCGGGCATCAGTTTGAAGAACTAATCTCTTCAAGATTTCTCAATCCCGAT 540
Db 481 GAGAAAGCTAAGCGGGCATCAGTTTGAAGAACTAATCTCTTCAAGATTTCTCAATCCCGAT 540
Qy 541 GAAAGAGTGAAGTCCCTCCGCTCCGCTCCGACGAGAGCCAGCTGGGGACCACTTCCCGG 600
Db 541 GAAAGAGTGAAGTCCCTCCGCTCCGCTCCGACGAGAGCCAGCTGGGGACCACTTCCCGG 600
Qy 601 GAGCAAGGCGCAGCCCTGCGGGCACTTCTCAAGCGCAGACAGATTGATTTCCGCTCGCG 660
Db 601 GAGCAAGGCGCAGCCCTGCGGGCACTTCTCAAGCGCAGACAGATTGATTTCCGCTCGCG 660
Qy 661 ATCTGTGTCCTCCACCCAGTTTGTGTGTCATCATGGAAGAGAGGCTTGACATTAAG 720
Db 661 ATCTGTGTCCTCCACCCAGTTTGTGTGTCATCATGGAAGAGAGGCTTGACATTAAG 720
Qy 721 AACATATCTAAGAGAGCCAGTCCGGGTAGATATCATAGAAAAGAACTCTGAGGCT 780
Db 721 AACATATCTAAGAGAGCCAGTCCGGGTAGATATCATAGAAAAGAACTCTGAGGCT 780
Qy 781 GCAGAGAAGCTGTCAACATCATGTCACCCGAGAGGGGACTTCTGAAGCATGCGCATG 840
Db 781 GCAGAGAAGCTGTCAACATCATGTCACCCGAGAGGGGACTTCTGAAGCATGCGCATG 840
Qy 841 ATTCTTGAATCATGCAAGAAAGGCGATGAGAACCCAACTAGCCGAAAGATTTCTCTG 900
Db 841 ATTCTTGAATCATGCAAGAAAGGCGATGAGAACCCAACTAGCCGAAAGATTTCTCTG 900
Qy 901 AAAATCTGGGCAACAATGGCTTGTTGGAAGACTGATTGAAAAAGAGGCAAAAATTTG 960
Db 901 AAAATCTGGGCAACAATGGCTTGTTGGAAGACTGATTGAAAAAGAGGCAAAAATTTG 960
Qy 961 AAGAAAATTGACATGAAACAGGGACCCAGATTAACAATCTCATCTTTGACGATTTGAGC 1020
Db 961 AAGAAAATTGACATGAAACAGGGACCCAGATTAACAATCTCATCTTTGACGATTTGAGC 1020
Qy 1021 ATATTAACACCCGGAAGAACATCATCTGTGAAGGGCAAGTTGAGGCTGTGCGAGTGT 1080
Db 1021 ATATTAACACCCGGAAGAACATCATCTGTGAAGGGCAAGTTGAGGCTGTGCGAGTGT 1080
Qy 1081 GAATATGAGATTATGAAGAAGCTGCGTGAAGGCTTTGAAAATGATATGCTGCTTTAAC 1140
Db 1081 GAATATGAGATTATGAAGAAGCTGCGTGAAGGCTTTGAAAATGATATGCTGCTTTAAC 1140
Qy 1141 ACCCACTCCGGAATCTTCCAGCTGTACCCCATCAACAGTTTGCCGTTCCGCAT 1200
Db 1141 ACCCACTCCGGAATCTTCCAGCTGTACCCCATCAACAGTTTGCCGTTCCGCAT 1200

Db 1141 ACCCACTCCGGAATCTTCCAGCTGTACCCCATCAACAGTTTGCCGTTCCGCAT 1200
Qy 1201 CATCACTCTTATCCAGAGCAGAGATTGTGAATCTCTTCATCCCAACCAAGCTGTGGC 1260
Db 1201 CATCACTCTTATCCAGAGCAGAGATTGTGAATCTCTTCATCCCAACCAAGCTGTGGC 1260
Qy 1261 GGCATCATCGGGAAGAGGGGGCACATCAAAACGCTGGGGAGATTGCGCGAGCCTCT 1320
Db 1261 GGCATCATCGGGAAGAGGGGGCACATCAAAACGCTGGGGAGATTGCGCGAGCCTCT 1320
Qy 1321 ATCAAGATTGGCCCTGGGGAAGGCGCAAGCGTACGCGAAAGGATGGTCATCATCCGGG 1380
Db 1321 ATCAAGATTGGCCCTGGGGAAGGCGCAAGCGTACGCGAAAGGATGGTCATCATCCGGG 1380
Qy 1381 CCAACGGAAGCCAGTTCAAGGCCAGGGACCGATCTTTGGGAAACTGAAGAAGAAAAC 1440
Db 1381 CCAACGGAAGCCAGTTCAAGGCCAGGGACCGATCTTTGGGAAACTGAAGAAGAAAAC 1440
Qy 1441 TTCTTTAAACCCCAAGAAAGAGTGAAGCTGGAAGGCTATTCAGATGCCCTTCCACA 1500
Db 1441 TTCTTTAAACCCCAAGAAAGAGTGAAGCTGGAAGGCTATTCAGATGCCCTTCCACA 1500
Qy 1501 GCTGCGCGGGTGAATGGCAAGGTTGGCAAGACCGTGAACGAACCTCAGAACTTAAACAGT 1560
Db 1501 GCTGCGCGGGTGAATGGCAAGGTTGGCAAGACCGTGAACGAACCTCAGAACTTAAACAGT 1560
Qy 1561 GCAGAAATCATGTCCTCTGTGACCAAAACGCCAGATGAATGAAGAAAGTATGCTCAGA 1620
Db 1561 GCAGAAATCATGTCCTCTGTGACCAAAACGCCAGATGAATGAAGAAAGTATGCTCAGA 1620
Qy 1621 ATTATCGGGCACTTTCTTGCTAGCCAGACTGACAGCGGAATCAGAGGAAATTTGACAA 1680
Db 1621 ATTATCGGGCACTTTCTTGCTAGCCAGACTGACAGCGGAATCAGAGGAAATTTGACAA 1680
Qy 1681 CAGGTGAAGCAGCAGAGAGCAAGAAATACCTCAGGGAGTGCCTCAACAGCGCAAGTGA 1740
Db 1681 CAGGTGAAGCAGCAGAGAGCAAGAAATACCTCAGGGAGTGCCTCAACAGCGCAAGTGA 1740
Qy 1741 GGCCTCCACAGCACCAGCAAAACAAACGAGTGAATGATGACCTTCCAAACCTGACAGAA 1800
Db 1741 GGCCTCCACAGCACCAGCAAAACAAACGAGTGAATGATGACCTTCCAAACCTGACAGAA 1800
Qy 1801 TGAAGCAAAACGACGACCCAGATCGGGAGCAAAACCAAAAGACATCTGAGAAATGAGAA 1860
Db 1801 TGAAGCAAAACGACGACCCAGATCGGGAGCAAAACCAAAAGACATCTGAGAAATGAGAA 1860
Qy 1861 GTCTGCGAGCGCGCAGAGGACTCTGCGAGGCTTGAGAACCCCAAGGCGCAGAGAGG 1920
Db 1861 GTCTGCGAGCGCGCAGAGGACTCTGCGAGGCTTGAGAACCCCAAGGCGCAGAGAGG 1920
Qy 1921 GCGGGAGAGGTCAAGGCTTTGCGAAGAACCAACGAGCCCTCTCCGCCCCCAGGGC 1980
Db 1921 GCGGGAGAGGTCAAGGCTTTGCGAAGAACCAACGAGCCCTCTCCGCCCCCAGGGC 1980
Qy 1981 TTCTGAGGCTTCAAGCATCACTTCAACATCACTCGATCTCTCTGAACTCCACGA 2040
Db 1981 TTCTGAGGCTTCAAGCATCACTTCAACATCACTCGATCTCTCTGAACTCCACGA 2040
Qy 2041 CGCTATCCCTTTAGTTGAATCAATAGGTGACGTTCAAAGCCAGCAAAATGAC 2100
Db 2041 CGCTATCCCTTTAGTTGAATCAATAGGTGACGTTCAAAGCCAGCAAAATGAC 2100
Qy 2101 ACCCTTTTCTGTGGCAATGCTGTCTGTACATGTGTACATATTAAGAAAGGAAATG 2160
Db 2101 ACCCTTTTCTGTGGCAATGCTGTCTGTACATGTGTACATATTAAGAAAGGAAATG 2160
Qy 2161 TTAAGATATGTCCTGTGGTTTACAGAGGTGCTCAGCGGTAATATATTTAGAAAT 2220
Db 2161 TTAAGATATGTCCTGTGGTTTACAGAGGTGCTCAGCGGTAATATATTTAGAAAT 2220
Qy 2221 AATATATCAATATCACTAATCTCAATTTTAACTATTAATTTTCTTT 2280
Db 2221 AATATATCAATATCACTAATCTCAATTTTAACTATTAATTTTCTTT 2280
```

QY	2281	TTAAAGAGAAACAGAGCTTTTCTACACTTTAAAGATTAAGTCTTTGGAGGCTCTACGG	2340
Db	2281	TTAAAGAGAAACAGAGCTTTTCTACACTTTAAAGATTAAGTCTTTGGAGGCTCTACGG	2340
QY	2341	TGTAGAGAGAGACTTTGAGGGCACCACCAAAATTCAACCCAGAGGAAATCTGTGGGA	2400
Db	2341	TGTAGAGAGAGACTTTGAGGGCACCACCAAAATTCACCCAGAGGAAATCTGTGGGA	2400
QY	2401	AGGACACTCAGCGGCACTTGAGATCACTGTGTATGTCAACAGAAAGGATACCGTCTCT	2460
Db	2401	AGGACACTCAGCGGCACTTGAGATCACTGTGTATGTCAACAGAAAGGATACCGTCTCT	2460
QY	2461	TGAAGAGGAAACTCTGTACTCTTCACATCCGCTGTCTAGTCAATACCCATTTCTCTTTC	2520
Db	2461	TGAAGAGGAAACTCTGTACTCTTCACATCCGCTGTCTAGTCAATACCCATTTCTCTTTC	2520
QY	2521	TTCAAGGTTTTAAACTGGTTTTTTTGACATACGTCAATTAATTCGTCTGTCTCTGGT	2580
Db	2521	TTCAAGGTTTTAAACTGGTTTTTTTGACATACGTCAATTAATTCGTCTGTCTCTGGT	2580
QY	2581	TATCTCTCCCTCCCTCCCTCCCTCCCTCTTCTCATCTCATCTGATTTTGAATTTCTCAT	2640
Db	2581	TATCTCTCCCTCCCTCCCTCCCTCCCTCTTCTCATCTCATCTGATTTTGAATTTCTCAT	2640
QY	2641	CCCTCCCATCTCAATCCCGTATCTTAGAGACCCCGCCCGCCCGCAGGCAAGAGTCTCTGA	2700
Db	2641	CCCTCCCATCTCAATCCCGTATCTTAGAGACCCCGCCCGCCCGCAGGCAAGAGTCTCTGA	2700
QY	2701	GTATCATCATCACAGAAAGAAAGCAAAAGCAAAACACAAACAGGCTCAACTTACACTT	2760
Db	2701	GTATCATCATCACAGAAAGAAAGCAAAAGCAAAACACAAACAGGCTCAACTTACACTT	2760
QY	2761	GGTATCTCAAAAGAAACAAGTCAATGGTACTTGTCTTCCACGTTTTGGAAGAGAAACA	2820
Db	2761	GGTATCTCAAAAGAAACAAGTCAATGGTACTTGTCTTCCACGTTTTGGAAGAGAAACA	2820
QY	2821	GGAACCCACCAACCAACCAATCAACCAACCAAGAAAAATTCACAAATGAAGAAATGT	2880
Db	2821	GGAACCCACCAACCAACCAATCAACCAACCAAGAAAAATTCACAAATGAAGAAATGT	2880
QY	2881	ATTTTGTCTTTTGGATTTTGGTGTATAGCAATCAATTTACAGCAAAATGATTCCTTTC	2940
Db	2881	ATTTTGTCTTTTGGATTTTGGTGTATAGCAATCAATTTACAGCAAAATGATTCCTTTC	2940
QY	2941	TTTTAAAAAATAATGTGAGGAAAGTAAATTTACAAAGTTGTGTCGCCACGAGCGTT	3000
Db	2941	TTTTAAAAAATAATGTGAGGAAAGTAAATTTACAAAGTTGTGTCGCCACGAGCGTT	3000
QY	3001	AAATTCACAGATTTTAAACGAGAAAAACACACAGAGAGCTACTCAGAGTGTTTTA	3060
Db	3001	AAATTCACAGATTTTAAACGAGAAAAACACACAGAGAGCTACTCAGAGTGTTTTA	3060
QY	3061	CTCAGACACTTGCTGTGTGTTCCTTAAGAGATTTGTAAAGCTGATAGTTGAGCAT	3120
Db	3061	CTCAGACACTTGCTGTGTGTTCCTTAAGAGATTTGTAAAGCTGATAGTTGAGCAT	3120
QY	3121	TTTTTATTTTTTAAATAAAGAGTTGGAATAAATAATAGATATCACTGCACGCTTG	3180
Db	3121	TTTTTATTTTTTAAATAAAGAGTTGGAATAAATAATAGATATCACTGCACGCTTG	3180
QY	3181	GAGAGGTGACAGTCCAGAGTGCACAAGCTGTTCTGAATTTGTCTTCCTGACGCAAGAA	3240
Db	3181	GAGAGGTGACAGTCCAGAGTGCACAAGCTGTTCTGAATTTGTCTTCCTGACGCAAGAA	3240
QY	3241	CCAAATATGAGCTTTCTTTGGACAAACCTTGAAAAATGTTAATTT	3283
Db	3241	CCAAATATGAGCTTTCTTTGGACAAACCTTGAAAAATGTTAATTT	3283

```

1 PUBLICATION NO. US20030165831A1
2 GENERAL INFORMATION:
3 APPLICANT: Lee, John
4 APPLICANT: Thompson, Pamela
5 APPLICANT: Lillie, James
6 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
7 TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
8 TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
9 FILE REFERENCE: MRI-006B
10 CURRENT APPLICATION NUMBER: US/09/814,353
11 CURRENT FILING DATE: 2001-03-21
12 PRIOR APPLICATION NUMBER: US 60/191,031
13 PRIOR FILING DATE: 2000-03-21
14 PRIOR APPLICATION NUMBER: US 60/207,124
15 PRIOR FILING DATE: 2000-05-25
16 PRIOR APPLICATION NUMBER: US 60/211,940
17 PRIOR FILING DATE: 2000-06-15
18 PRIOR APPLICATION NUMBER: US 60/216,820
19 PRIOR FILING DATE: 2000-07-07
20 PRIOR APPLICATION NUMBER: US 60/220,661
21 PRIOR FILING DATE: 2000-07-25
22 PRIOR APPLICATION NUMBER: US 60/257,672
23 PRIOR FILING DATE: 2000-12-21
24 NUMBER OF SEQ ID NOS: 22037
25 SOFTWARE: FastSeq for Windows Version 4.0
26 SEQ ID NO 20241
27 LENGTH: 3905
28 TYPE: DNA
29 ORGANISM: Homo sapiens
30 FEATURE:
31 NAME/KEY: misc_feature
32 LOCATION: 1, 3897, 3898, 3899, 3900, 3901, 3902, 3903, 3904, 3905
33 OTHER INFORMATION: n = A,T,C or G
34 IS-09-814-353-20241

```

[illegible]

Db 849 GAAGCTTAGCGGGGCTAGTTTGAACTACTCTTCAAGATTTCCTACATCCCGATGA 908
QY 543 AGAGGTAGCTCCCTTTCGCCCCCTCAGCGAGCCAGCGTGGGGACCACTTTCCCGGGA 602
Db 909 AGAGGTAGCTCCCTTTCGCCCCCTCAGCGAGCCAGCGTGGGGACCACTTTCCCGGGA 968
QY 603 GCAGGCGCAGCCCTTGGGGGCACTTCTCAGGCCAGAGATGATTTCCCGTGGCAT 662
Db 969 GCAAGGCCAGCCCTTGGGGGCACTTCTCAGGCCAGAGATGATTTCCCGTGGCAT 1028
QY 663 CTTGATCCCGCAGCTTGTGTGTGCATCATCGAAGAGAGGAGGCTTGACATTAAGAA 722
Db 1029 CTTGATCCCGCAGCTTGTGTGTGCATCATCGAAGAGAGGAGGCTTGACATTAAGAA 1088
QY 723 CATCATTAAGCAGACCCAGTCCCGGGTATGATTCATTAAGAAAGAACTCTGAGCTGC 782
Db 1089 CATCATTAAGCAGACCCAGTCCCGGGTATGATTCATTAAGAAAGAACTCTGAGCTGC 1148
QY 783 AGAGAGGCTGTGCATCCATCCAGTCCAGAGGGGAGCTTCTGAAGCATCCGCATGAT 842
Db 1149 AGAGAGGCTGTGCATCCATCCAGTCCAGAGGGGAGCTTCTGAAGCATCCGCATGAT 1208
QY 843 TCTTGAATCATGCGAAGAGGCGAGATGAGACCAACTAGCCGAAAGATTTCTTGAA 902
Db 1209 TCTTGAATCATGCGAAGAGGCGAGATGAGACCAACTAGCCGAAAGATTTCTTGAA 1268
QY 903 AATCTTGGCACAATGCGCTTGTGTGTGAAAGCTGATTTGAAAAAGAGCGAAGAAATTGAA 962
Db 1269 AATCTTGGCACAATGCGCTTGTGTGTGAAAGCTGATTTGAAAAAGAGCGAAGAAATTGAA 1328
QY 963 GAAATTTAAACATGAAACAGGGGACCAAGTAACTCATCTTTGAGATTTGAGCAT 1022
Db 1329 GAAATTTAAACATGAAACAGGGGACCAAGTAACTCATCTTTGAGATTTGAGCAT 1388
QY 1023 ATACACCCGGAAGAACCATCATCTGTGAAGGGCAAGTTGAGGCTGTGCCAGTGCTGA 1082
Db 1389 ATACACCCGGAAGAACCATCATCTGTGAAGGGCAAGTTGAGGCTGTGCCAGTGCTGA 1448
QY 1083 GATAGAGATTATGAAGAGCTGCTGTGAGGCTTTGAAATGATATGCTGGCTGTTAAGAC 1142
Db 1449 GATAGAGATTATGAAGAGCTGCTGTGAGGCTTTGAAATGATATGCTGGCTGTTAAGAC 1508
QY 1143 CCAGCTCCGGAATCTTCCAGGCTGTGAGGCTTCCAGTCCAGAGTTGGCCCGTCCCGCATCA 1202
Db 1509 CCAGCTCCGGAATCTTCCAGGCTGTGAGGCTTCCAGTCCAGAGTTGGCCCGTCCCGCATCA 1568
QY 1203 TCACTCTTATCCAGAGCAGAGATTTGTGAATCTTTTATCCAACTCAGGCTGTGGGCGC 1262
Db 1569 TCACTCTTATCCAGAGCAGAGATTTGTGAATCTTTTATCCAACTCAGGCTGTGGGCGC 1628
QY 1263 CATCATCGGAGAGAGGGGGCACAATCAAAAGCTGGCGAGATTTGGCGGAGACCTTAT 1322
Db 1629 CATCATCGGAGAGAGGGGGCACAATCAAAAGCTGGCGAGATTTGGCGGAGACCTTAT 1688
QY 1323 CAAGATTGCCCCCGGGAAGGCCAGAGCTCAGCGAAAGAGATGTCATCATCACTCCGAGC 1382
Db 1689 CAAGATTGCCCCCGGGAAGGCCAGAGCTCAGCGAAAGAGATGTCATCATCACTCCGAGC 1748
QY 1383 ACCGGAAGCCAGATTCAAGGCCAGAGCGGATCTTTGGGAAACTGAAAGAGAAACTT 1442
Db 1749 ACCGGAAGCCAGATTCAAGGCCAGAGCGGATCTTTGGGAAACTGAAAGAGAAACTT 1808
QY 1443 CTTTAAACCCCAAGAGAGATGAAAGCTGGAAGGCGCATATCAAGATGCGCTTCCACAGC 1502
Db 1809 CTTTAAACCCCAAGAGAGATGAAAGCTGGAAGGCGCATATCAAGATGCGCTTCCACAGC 1868
QY 1503 TGGCCGGGTGATTGGCAAGGTGCGCAGAGCGTGAAGCAACTGCAAGACTTAACTCACTGC 1562
Db 1869 TGGCCGGGTGATTGGCAAGGTGCGCAGAGCGTGAAGCAACTGCAAGACTTAACTCACTGC 1928
QY 1563 AGAAGTCATGTCCTGTGTGACCAAGCGCATGAAGAAATGAGGAAGTATGTCAGAAAT 1622
Db 1929 AGAAGTCATGTCCTGTGTGACCAAGCGCATGAAGAAATGAGGAAGTATGTCAGAAAT 1988

QY 1623 TATGGGCACTTCTTTGCTAAGCCAGACTGCAAGCGCAAGTCAAGGAAATTTGACACA 1682
Db 1989 TATGGGCACTTCTTTGCTAAGCCAGACTGCAAGCGCAAGTCAAGGAAATTTGACACA 2048
QY 1683 GGTGAAGCAGCAGAGAGCAAAATTCCTTCAGGGAGTGCCCTCAAGCGCAGCAAGTAGG 1742
Db 2049 GGTGAAGCAGCAGAGAGCAAAATTCCTTCAGGGAGTGCCCTCAAGCGCAGCAAGTAGG 2108
QY 1743 CTCCCAAGGCAACAGCAAAACAACGATGATTAAGCTTTCACACCTGACAGAAATG 1802
Db 2109 CTCCCAAGGCAACAGCAAAACAACGATGATTAAGCTTTCACACCTGACAGAAATG 2168
QY 1803 AGACCAACGCAAGCAGCAGATGGAGCAACCAACCAACCAATCTGAGATGAGAGT 1862
Db 2169 AGACCAACGCAAGCAGCAGATGGAGCAACCAACCAACCAATCTGAGATGAGAGT 2228
QY 1863 CTGGGAGGCGGCCAGAGGACTCTGCGGAGGCTCGAGAACCCAGAGGSCCAGAGAGGGC 1922
Db 2229 CTGGGAGGCGGCCAGAGGACTCTGCGGAGGCTCGAGAACCCAGAGGSCCAGAGAGGGC 2288
QY 1923 GGGGAAGGTCAAGCAGTTTGCAGAACCAACGAGCCCGCTCCGCCCCCAGGGCTT 1982
Db 2289 GGGGAAGGTCAAGCAGTTTGCAGAACCAACGAGCCCGCTCCGCCCCCAGGGCTT 2348
QY 1983 CTGAGGCTTCAGCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 2042
Db 2349 CTGAGGCTTCAGCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 2408
QY 2043 CTATCCCTTTAGTTGAATCAATGAGTGAACGTTCAAGAGCAAGCAAAATGACAC 2102
Db 2409 CTATCCCTTTAGTTGAATCAATGAGTGAACGTTCAAGAGCAAGCAAAATGACAC 2468
QY 2103 CTTTTCCTGGCAAAATGCTCTGTACATGTGTACATGTGTACATGTGTACATGTGTACAT 2162
Db 2469 CTTTTCCTGGCAAAATGCTCTGTACATGTGTACATGTGTACATGTGTACATGTGTACAT 2528
QY 2163 AAGATATGCGCTGTGGTTTCAAGGGTGCCTGAGGGGTATATTTAAGAAATA 2222
Db 2529 AAGATATGCGCTGTGGTTTCAAGGGTGCCTGAGGGGTATATTTAAGAAATA 2588
QY 2223 TATATCAAAATCTCAACTCACTCAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 2282
Db 2589 TATATCAAAATCTCAACTCACTCAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 2648
QY 2283 AAAGAGAAAGCAGGCTTTTCTAGACTTAAAGAAATAGCTTTTGGAGGTCTCAGGTTG 2342
Db 2649 AAAGAGAAAGCAGGCTTTTCTAGACTTAAAGAAATAGCTTTTGGAGGTCTCAGGTTG 2708
QY 2343 TAGAGGAGCTTTTGAAGCCAGCCGCAAAATTCACCCAGAGGAAATCTGTCGGAAAG 2402
Db 2709 TAGAGGAGCTTTTGAAGCCAGCCGCAAAATTCACCCAGAGGAAATCTGTCGGAAAG 2768
QY 2403 GACACTCAGGCGAGTTCTGGATCACTGTGTATGTCAAGAGAGGATACCGTCTCCTTG 2462
Db 2769 GACACTCAGGCGAGTTCTGGATCACTGTGTATGTCAAGAGAGGATACCGTCTCCTTG 2828
QY 2463 AAGAGAAACTCTGTCACTCTCATGCTGTGTAGCTCATACCCATTTCTCTTGGCTT 2522
Db 2829 AAGAGAAACTCTGTCACTCTCATGCTGTGTAGCTCATACCCATTTCTCTTGGCTT 2888
QY 2523 CACAGTTTAAACTGTTTGTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2582
Db 2889 CACAGTTTAAACTGTTTGTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2948
QY 2583 TCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 2642
Db 2949 TCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3008
QY 2643 CTCATCTCAATCCCGATCTAGCAGACCCCGCCCGCCAGGCAAGAGAGTGTCTGAGT 2702
Db 3009 CTCATCTCAATCCCGATCTAGCAGACCCCGCCCGCCAGGCAAGAGAGTGTCTGAGT 3067

QY 2703 ATCATATACACAAAGGAAACAAAGCAACACAAACAGCTCAACTTACCTTGG 2762
 Db 3068 ATCATATACACAAAGGAAACAAAGCAACACAAACAGCTCAACTTACCTTGG 3127
 QY 2763 TTACTCAAAAGCAAGGTCATGTAATCTGTCTTACGCTTTGGAGAGAAACAG 2822
 Db 3128 TTACTCAAAAGCAAGGTCATGTAATCTGTCTTACGCTTTGGAGAGAAACAG 3187
 QY 2823 AACCCACAAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAAT 2882
 Db 3188 AACCCACAAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAAT 3247
 QY 2883 TTGTCTTTTTCATTTTGTGTATTAAGCCATCAATATTTAGCAAAATGATTCCTTCT 2942
 Db 3248 TTGTCTTTTTCATTTTGTGTATTAAGCCATCAATATTTAGCAAAATGATTCCTTCT 3307
 QY 2943 T-AAAAAAAAAATGTGAGGAAAGTAGAAATTTACCAAGGTTGTGGCCAGGCGCTTA 3001
 Db 3308 TAAAAAAAAAATGTGAGGAAAGTAGAAATTTACCAAGGTTGTGGCCAGGCGCTTA 3367
 QY 3002 AATTCAGAGATTTTAAACAGAGAAACCAACAGAGAGAGCTTCTCAGGTGTTTTAC 3061
 Db 3368 AATTCAGAGATTTTAAACAGAGAAACCAACAGAGAGAGCTTCTCAGGTGTTTTAC 3427
 QY 3062 CTCAGACCTTGTCTTGTGTGTTTCCCTTAGAGATTTTGAAGCTGATGTTGGAGATT 3121
 Db 3428 CTCAGACCTTGTCTTGTGTGTTTCCCTTAGAGATTTTGAAGCTGATGTTGGAGATT 3487
 QY 3122 TTTTATTTTATTAATAAATGAGTTGAAATTAAGATATCAATGCGCAGCTGG 3181
 Db 3488 TTTTATTTTATTAATAAATGAGTTGAAATTAAGATATCAATGCGCAGCTGG 3547
 QY 3182 AGAAGGTGACAGTCCAGTGTGCAAGCTGTTCTGAATGTCTTCCGCTAGCCAGAAC 3241
 Db 3548 AGAAGGTGACAGTCCAGTGTGCAAGCTGTTCTGAATGTCTTCCGCTAGCCAGAAC 3607
 QY 3242 CNAATGCGCTTCTTTTGAACAAACCTTGAATGTTTATTT 3283
 Db 3608 CNAATGCGCTTCTTTTGAACAAACCTTGAATGTTTATTT 3649

RESULT 3
 US-10-097-340-146
 ; Sequence 146, Application US/10097340
 ; Publication No. US20030087250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: JOHN MONAHAN
 ; APPLICANT: Manjula GANNAVAPU
 ; APPLICANT: Sebastian HOERSCHE
 ; APPLICANT: Shubhangi KAMATKAR
 ; APPLICANT: Steve G. KOVATS
 ; APPLICANT: Rachel E. MEYERS
 ; APPLICANT: Michael MORRISSEY
 ; APPLICANT: Peter OLANDT
 ; APPLICANT: Ami SEN
 ; APPLICANT: Peter VEIBY
 ; APPLICANT: Gordon B. MILLS
 ; APPLICANT: Robert C. BAST, Jr.
 ; APPLICANT: Karen LU
 ; APPLICANT: Rosemarie SCHMANDT
 ; APPLICANT: Xumei ZHAO
 ; APPLICANT: Karen GLATT
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
 ; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
 ; FILE REFERENCE: MRI-030
 ; CURRENT APPLICATION NUMBER: US/10/097,340
 ; PRIOR FILING DATE: 2002-03-14
 ; PRIOR APPLICATION NUMBER: 60/276,025
 ; PRIOR FILING DATE: 2001-03-14
 ; PRIOR APPLICATION NUMBER: 60/325,149
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 60/276,026
 ; PRIOR FILING DATE: 2001-03-14

;; PRIOR APPLICATION NUMBER: 60/324,967
 ;; PRIOR FILING DATE: 2001/09/26
 ;; PRIOR APPLICATION NUMBER: 60/311,732
 ;; PRIOR FILING DATE: 2001-08-10
 ;; PRIOR APPLICATION NUMBER: 60/325,102
 ;; PRIOR FILING DATE: 2001-09-26
 ;; PRIOR APPLICATION NUMBER: 60/323,580
 ;; PRIOR FILING DATE: 2001-09-19
 ;; NUMBER OF SEQ ID NOS: 363
 ;; SOFTWARE: FASTSEQ for Windows Version 4.0
 ;; SEQ ID NO 146
 ;; LENGTH: 3667
 ;; TYPE: DNA
 ;; ORGANISM: Homo sapiens
 US-10-097-340-146

Query Match 99.1%; Score 3253.8; DB 15; Length 3667;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 3277; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 3 CAGCGAGAGAGCGAGAGAGCGCGGTACCGCGCGCGGAGCGCGCGCGCTCTCGGGGAA 62
 Db 369 CAACGAGAGAGCGAGAGAGCGCGGTACCGCGCGCGGAGCGCGCGCGCTCTCGGGAA 428
 QY 63 GAGACGATGATGAACAAGCTTTTACATGCGAACTTGAACCGCGCGGTACCGCGGAGA 122
 Db 429 GAGACGATGATGAACAAGCTTTTACATGCGAACTTGAACCGCGCGGTACCGCGGAGA 488
 QY 123 CTTCCCGGACCTCTTGTGGGAGCAGAGAGCTGCGCGCGGAGAGAGTCTTGTGAAGTC 182
 Db 489 CTTCCCGGACCTCTTGTGGGAGCAGAGAGCTGCGCGCGGAGAGAGTCTTGTGAAGTC 548
 QY 183 CCGCTACGCTTGTGACTACCCCGACAGAACTGGGCGCATCGCGCATCGAGACCT 242
 Db 549 CCGCTACGCTTGTGACTACCCCGACAGAACTGGGCGCATCGCGCATCGAGACCT 608
 QY 243 CTGCGGTAAAGTGAATGATGAGAAATCATGAAAGTTGATTTCTCACTTTTAAAA 302
 Db 609 CTGCGGTAAAGTGAATGATGAGAAATCATGAAAGTTGATTTCTCACTTTTAAAA 668
 QY 303 GCTAAGAGAGAGAAATGATGAGAAATCATGAAATCTTCTCACTGAGTGGAGGTT 362
 Db 669 GCTAAGAGAGAGAAATGATGAGAAATCATGAAATCTTCTCACTGAGTGGAGGTT 728
 QY 363 GGATGACCTTTTGTCAATATGAGCAAGTGAAGTGTGAAACAAGTCAACAGAC 422
 Db 729 GGATGACCTTTTGTCAATATGAGCAAGTGAAGTGTGAAACAAGTCAACAGAC 788
 QY 423 AGAAACCGCGCTTGTCAAGTCAATATGCAACAAGAAAGAAATTAAGCCATGGA 482
 Db 789 AGAAACCGCGCTTGTCAAGTCAATATGCAACAAGAAAGAAATTAAGCCATGGA 848
 QY 483 GAAGCTAAGCGGCGCATCAGTTGAGAACTACTCTTCAAGTTTCTTCAATCCCGATGA 542
 Db 849 GAAGCTAAGCGGCGCATCAGTTGAGAACTACTCTTCAAGTTTCTTCAATCCCGATGA 908
 QY 543 AGAGTGAAGCTCCCTTTCAGGAGAGCGCGTGGGAGCACTTCCCGGGA 602
 Db 909 AGAGTGAAGCTCCCTTTCAGGAGAGCGCGTGGGAGCACTTCCCGGGA 968
 QY 603 GCAAGGCGACGCGCTTGGGGGAGCACTTTCAGGCGAGAGATTTTCCCGTGGAGT 662
 Db 969 GCAAGGCGACGCGCTTGGGGGAGCACTTTCAGGCGAGAGATTTTCCCGTGGAGT 1028
 QY 663 CCGTGTCCCGACCGAGTTTGTGGGCGCATCGGAAAGAGGCGTTGACCTTAAGAA 722
 Db 1029 CCGTGTCCCGACCGAGTTTGTGGGCGCATCGGAAAGAGGCGTTGACCTTAAGAA 1088
 QY 723 CATCATAGAGAGACCGAGTCCCGGAGATATTCATAGAAAGAACTTGGAGCTGC 782
 Db 1089 CATCATAGAGAGACCGAGTCCCGGAGATATTCATAGAAAGAACTTGGAGCTGC 1148
 QY 783 AGAAGAGCTGTACCATTCATGCCACCCAGAGGGAGCTTCTGAAGCATGCGCATGAT 842

1149 AGAAGACCTGTCACTCCATGCGACCGAGGAGCTTCTAGAGCTCCGATCAT 1208
843 TCTTGAATCATGAG 902
1209 TCTTGAATCATGAG 1268
903 AATCTGGGACACAAATGGCTTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962
1269 AATCTGGGACACAAATGGCTTGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1328
963 GAAATTTGAACATGAG 1022
1329 GAAATTTGAACATGAG 1388
1023 ATAGAACCCGAG 1082
1389 ATAGAACCCGAG 1448
1083 GATGAGAGATTTGAG 1142
1449 GATGAGAGATTTGAG 1508
1143 GATGAGAGATTTGAG 1202
1509 GATGAGAGATTTGAG 1568
1203 TCACTCTTATCCAG 1262
1569 TCACTCTTATCCAG 1628
1283 CATCATCCGAG 1322
1629 CATCATCCGAG 1688
1323 CAAGATGCGCCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1382
1689 CAAGATGCGCCCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1748
1383 ACCGGAAGCCGAG 1442
1749 ACCGGAAGCCGAG 1808
1443 CTTTAAACCCGAG 1502
1809 CTTTAAACCCGAG 1868
1503 TGGCCGAG 1562
1869 TGGCCGAG 1928
1563 AGAAGTCATGTCCTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1622
1929 AGAAGTCATGTCCTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1988
1623 TATCGGAG 1682
1989 TATCGGAG 2048
2049 GGTGAG 2108
1743 CTCCCAAGGAG 1802
2109 CTCCCAAGGAG 2168
1803 AGAGCAAAAGGAG 1862
2169 AGAGCAAAAGGAG 2228
1863 CTCCGAG 1922

2229 CTGCGAGAGCGGCGAG 2288
1923 GGGAGAGGTGAG 1982
2289 GGGAGAGGTGAG 2348
1983 CTGAGAGGTGAG 2042
2449 CTGAGAGGTGAG 2408
2043 CTATCCCTTTAGTTGAG 2102
2409 CTATCCCTTTAGTTGAG 2468
2103 CCTTTTCTGAG 2162
2469 CCTTTTCTGAG 2528
2163 AAGATATGTCCTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2222
2529 AAGATATGTCCTGTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2588
2223 TATATCAAAATGAG 2282
2589 TATATCAAAATGAG 2648
2283 AAAG 2342
2649 AAAG 2708
2343 TAG 2402
2709 TAG 2768
2403 GAGACCTCAGGAG 2462
2769 GAGACCTCAGGAG 2828
2463 AAG 2522
2829 AAG 2888
2523 CAG 2582
2889 CAG 2948
2583 TGTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 2642
2949 TGTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 3008
2643 CTCCATCTCAATCCGATATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2702
3009 CTCCATCTCAATCCGATATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3067
2703 ATGACATCAACAAAAG 2762
3068 ATGACATCAACAAAAG 3127
2763 TTATCAAAAAG 2822
3128 TTATCAAAAAG 3187
2823 AACCCCAAAAG 2882
3188 AACCCCAAAAG 3247
2883 TTGTCTTTTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 2942
3248 TTGTCTTTTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3307
2943 T-AAAAAAG 3001
3308 TAAAAAAG 3367


```

Db 1809 CTTTAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAAGTGGCCCTTCCACAGC 1868
QY 1503 TGGCCGGGTGATTTGGCAAAAGGTGGCAAGACCTGTAAACGAACTTAAACCAAGTGC 1562
Db 1869 TGGCCGGGTGATTTGGCAAAAGGTGGCAAGACCTGTAAACGAACTTAAACCAAGTGC 1928
QY 1563 AGAAGTCACTGTGCTCTGTGACCAAAAGCCAGATGAAAATGAGGAAGTGTGCAAT 1622
Db 1929 AGAAGTCACTGTGCTCTGTGACCAAAAGCCAGATGAAAATGAGGAAGTGTGCAAT 1988
QY 1623 TATCGGGCACTCTTGTGTAGCAGACTGCAAGCGCAAGATCAAGGAATTTGTACAA 1682
Db 1989 TATCGGGCACTCTTGTGTAGCAGACTGCAAGCGCAAGATCAAGGAATTTGTACAA 2048
QY 1683 GGTGAAGCAGCAGAGCAAGAAATACCTTCAAGGAAGTGGCTTCAAGCGCAAGATGAG 1742
Db 2049 GGTGAAGCAGCAGAGCAAGAAATACCTTCAAGGAAGTGGCTTCAAGCGCAAGATGAG 2108
QY 1743 CTCCCAAGCAGCAGCAAGAAACAAAGGATGATGAGCCCTTCAAGCACTGACAGATG 1802
Db 2109 CTCCCAAGCAGCAGCAAGAAACAAAGGATGATGAGCCCTTCAAGCACTGACAGATG 2168
QY 1803 AGACCAAAAGCAGCAGCAGATCGAGCAAAACCAAGACCATGTAGGAATGAGAAAT 1862
Db 2169 AGACCAAAAGCAGCAGCAGATCGAGCAAAACCAAGACCATGTAGGAATGAGAAAT 2228
QY 1863 CTGCGGAGGCGGCGCAGGGAATCTTGCAGAGGCGCTTGAAGAACCCCAAGGGCGGAGGGGC 1922
Db 2229 CTGCGGAGGCGGCGCAGGGAATCTTGCAGAGGCGCTTGAAGAACCCCAAGGGCGGAGGGGC 2288
QY 1923 GGGGGAAGGTCAAGGAGTTTGCAGAGACGAGAGCGCGCTCCCGCGCCCGCAGGAGCTT 1982
Db 2289 GGGGGAAGGTCAAGGAGTTTGCAGAGACGAGAGCGCGCTCCCGCGCCCGCAGGAGCTT 2348
QY 2289 GGGGGAAGGTCAAGGAGTTTGCAGAGACGAGAGCGCGCTCCCGCGCCCGCAGGAGCTT 2348
Db 1983 CTGCAAGGCTTCAAGCAGATCCATCTCAATCACTCGGATCTCTCTGAACTCCCAAGAGC 2042
QY 2349 CTGCAAGGCTTCAAGCAGATCCATCTCAATCACTCGGATCTCTCTGAACTCCCAAGAGC 2408
Db 2043 CTATCCCTTTAGTGAATCAATATGATGAGAACTGTTCAAAGCCCAAGGAATGACAC 2102
QY 2409 CTATCCCTTTAGTGAATCAATATGATGAGAACTGTTCAAAGCCCAAGGAATGACAC 2468
Db 2103 CCTTTTCTGTGGCAAAATGCTCTGTATACATGATGTATCAATATTAGAAAGGGAAGTGT 2162
QY 2469 CCTTTTCTGTGGCAAAATGCTCTGTATACATGATGTATCAATATTAGAAAGGGAAGTGT 2528
Db 2163 AAGATATGTGGCTGTGGGTATACAGAGGTGCTGACGCGTAAATATTTTGAATAA 2222
QY 2523 AAGATATGTGGCTGTGGGTATACAGAGGTGCTGACGCGTAAATATTTTGAATAA 2588
Db 2223 TATATCAAAATCACTCACTCACTCAATTTTATCAATTTTATTTTCTTTT 2282
QY 2589 TATATCAAAATCACTCACTCACTCAATTTTATCAATTTTATTTTCTTTT 2648
Db 2283 AAAGAAAGCAAGGCTTTCTAGACTTAAAGATTAAGCTTTGGAGGCTCAAGGCTG 2342
QY 2649 AAAGAAAGCAAGGCTTTCTAGACTTAAAGATTAAGCTTTGGAGGCTCAAGGCTG 2708
Db 2343 TAGAGAGAGGCTTTGAGGCGCACCCGCAAAATTCACCGAGGGAATCTGTGGAG 2402
QY 2709 TAGAGAGAGGCTTTGAGGCGCACCCGCAAAATTCACCGAGGGAATCTGTGGAG 2768
Db 2403 GACACTCAAGGCTTGTGATCACTGTGTATGTCAACGAGAGGATACCGTCTCTG 2462
QY 2769 GACACTCAAGGCTTGTGATCACTGTGTATGTCAACGAGAGGATACCGTCTCTG 2828
Db 2463 AAGAGAGAACTCTGTCACTCTCAATGCTGTAGCTCATACACCACTTCTCTGCTT 2522
QY 2829 AAGAGAGAACTCTGTCACTCTCAATGCTGTAGCTCATACACCACTTCTCTGCTT 2888
Db 2523 CACAGGTTTAACTGGTTTGTGCAATGCTAATATTTCTGTCTCTCTGTTA 2582
QY 2889 CACAGGTTTAACTGGTTTGTGCAATGCTAATATTTCTGTCTCTCTGTTA 2948

```

```

QY 2583 TCTCTCCCTCCCTCCCTCCCTCCCTCTCTCTCATCTCCATCTTTTGAATTTCTCATCC 2642
Db 2949 TCTCTCCCTCCCTCCCTCCCTCCCTCTCTCTCATCTCCATCTTTTGAATTTCTCATCC 3008
QY 2643 CTCATCTCAATCCCGTATCTTACGACACCCCGCCCCCCCCCAGGCAAGAGCTCTGAGT 2702
Db 3009 CTCATCTCAATCCCGTATCTTACGCA--CCCCCCCCCAGGCAAGAGCTCTGAGT 3067
QY 2703 ATCAATCAACAAAGGAAACAAAGGAAACACAAACAGCCTCACTTACACTTGG 2762
Db 3068 ATCAATCAACAAAGGAAACAAAGGAAACACAAACAGCCTCACTTACACTTGG 3127
QY 2763 TTAATCAAAAGCAAGATCAATGTATCTTGTCTTACGCTTTTGAAGAGGAAACAGG 2822
Db 3128 TTAATCAAAAGCAAGATCAATGTATCTTGTCTTACGCTTTTGAAGAGGAAACAGG 3187
QY 2823 AACCAACCAACCAACCAATCAATCAACCAACCAAGAAAAATTCACAAATGAATGTAT 2882
Db 3188 AACCAACCAACCAACCAATCAATCAACCAACCAAGAAAAATTCACAAATGAATGTAT 3247
QY 2883 TTTGTCTTTTGAATTTTGTGTATPACCATCAATATTTACGAAATGATTCCTTCT 2942
Db 3248 TTTGTCTTTTGAATTTTGTGTATPACCATCAATATTTACGAAATGATTCCTTCT 3307
QY 2943 T-AAAAAAAAAATGTGAGAGAAAGTAGAAATTTACAAAGTTGTGGCCCAAGGCGTTA 3001
Db 3308 TAAAAAAAAAATGTGAGAGAAAGTAGAAATTTACAAAGTTGTGGCCCAAGGCGTTA 3367
QY 3002 AATTCAAGATTTTAAACGAGAAAAACACAAAGAGCTACTCAAGGTTTAAAC 3061
Db 3368 AATTCAAGATTTTAAACGAGAAAAACACAAAGAGCTACTCAAGGTTTAAAC 3427
QY 3062 CTCAGACCTTGTCTGTGTGTCTTCTTGTAGAGATTTTGAAGTGTAGAGAT 3121
Db 3428 CTCAGACCTTGTGTGTGTGTCTTCTTGTAGAGATTTTGAAGTGTAGAGAT 3487
QY 3122 TTTTATTTTAAATTAATAAATGAGTTGAAAAAATTAAGATTCATGCGAGCTTGG 3181
Db 3488 TTTTATTTTAAATTAATAAATGAGTTGAAAAAATTAAGATTCATGCGAGCTTGG 3547
QY 3182 AAGAGTGACATGCCAGTGTGCAACAGCTGTCTGAATGTCTTCGCTAGCAAGAG 3241
Db 3548 AAGAGTGACATGCCAGTGTGCAACAGCTGTCTGAATGTCTTCGCTAGCAAGAG 3507
QY 3242 CNAATAGGCTCTTTTGACAAACCTTGAAATGTTATTT 3283
Db 3608 CNAATAGGCTCTTTTGACAAACCTTGAAATGTTATTT 3649

```

```

RESULT 5
US-09-899-651-6
/ Sequence 6, Application US/09899651
/ Patent No. US20020111470A1
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Yao-Tseng
/ APPLICANT: Gure, Ali
/ APPLICANT: Tseng, Solam
/ APPLICANT: Stockert, Elisabeth
/ APPLICANT: Jager, Elke
/ APPLICANT: Knuth, Alexander
/ APPLICANT: Old, Lloyd J.
/ TITLE OF INVENTION: Isolated Nucleic Acid Molecules Encoding Cancer
/ TITLE OF INVENTION: Associated
/ TITLE OF INVENTION: Antigen, The Antigens Per Se, And Uses Thereof
/ FILE REFERENCE: LUD 5538
/ CURRENT APPLICATION NUMBER: US/09/899,651
/ PRIOR APPLICATION NUMBER: 2001-07-06
/ PRIOR FILING DATE: 1998-04-17
/ NUMBER OF SEQ ID NOS: 8
/ SEQ ID NO 6
/ LENGTH: 3412

```



```
Db 2041 CGAGGAGGCGGGAGAGCTCAGCCAGTTTCCGAAACCAACGAGCCCGCTCCGCC 2100
Qy 1972 CCCCAGGAGCTTCGAGGCTTCAGCCATCCACTTCACATCCACTCGGATCTCTCTGA 2031
Db 2101 CCCCAGGAGCTTCGAGGCTTCAGCCATCCACTTCACATCCACTCGGATCTCTCTGA 2160
Qy 2032 CTCCACAGAGCTATCCCTTTTATGTTGAATCACTAGTGAAGGTTTCAAAAGCAAGC 2091
Db 2161 CTCCACAGAGCTATCCCTTTTATGTTGAATCACTAGTGAAGGTTTCAAAAGCAAGC 2220
Qy 2092 AAAATGACACACCTTTTCTGTGCAAAATCGTCTGTATCATGTGTATGATATTTGAAA 2151
Db 2221 AAAATGACACACCTTTTCTGTGCAAAATCGTCTGTATCATGTGTATGATATTTGAAA 2280
Qy 2152 GGGAGATGTTAAGATATGTGCTGTGTGTTACACAGGAGCTTCGAGCGGTATATAT 2211
Db 2281 GGGAGATGTTAAGATATGTGCTGTGTGTTACACAGGAGCTTCGAGCGGTATATAT 2340
Qy 2212 TTTAAGAAATATATATCAATTAATCACTAATCTCAATTTTATATATATATATTTT 2271
Db 2341 TTTAAGAAATATATATCAATTAATCACTAATCTCAATTTTATATATATATTTT 2400
Qy 2272 TTTTCTTTTAAAGAGAGAGGCTTTTCTAGACTTTAAAGATAAAGTCTTTGGAG 2331
Db 2401 TTTTCTTTTAAAGAGAGAGGCTTTTCTAGACTTTAAAGATAAAGTCTTTGGAG 2460
Qy 2332 GTCTACGCTGTAGAGAGAGGCTTTGAGGCAACCGGACAAATTTACCAAGGGAAT 2391
Db 2461 GTCTACGCTGTAGAGAGAGGCTTTGAGGCAACCGGACAAATTTACCAAGGGAAT 2520
Qy 2392 CTCGTGGAAGAGACCTCAGGAGAGTCTGATCACTGTATGTCAACAGAGGATA 2451
Db 2521 CTCGTGGAAGAGACCTCAGGAGAGTCTGATCACTGTATGTCAACAGAGGATA 2580
Qy 2452 CCGTCTCTTGAAGAGAACTCTGTCACTCTCATGCTGTCTAGCTATACACCATT 2511
Db 2581 CCGTCTCTTGAAGAGAACTCTGTCACTCTCATGCTGTCTAGCTATACACCATT 2640
Qy 2512 TCTCTTGTTCACAGGTTTAACTGTGTTTGTGATCTGTATATATATCTCTGTCT 2571
Db 2641 TCTCTTGTTCACAGGTTTAACTGTGTTTGTGATCTGTATATATATCTCTGTCT 2700
Qy 2572 CTCTCTGTTCATCTCCCTCCCTCCCTCCCTCCCTCTCTCCATCTCCATCTTTTGA 2631
Db 2701 CTCTCTGTTCATCTCCCTCCCTCCCTCCCTCCCTCTCTCCATCTCTTTTGA 2760
Qy 2632 TTTCTCATCTCCCTCATCTCAATCCGTATCTAGCAGCCGCCGCCCGGCAAGCA 2691
Db 2761 TTTCTCATCTCCCTCATCTCAATCCGTATCTAGCAGCCGCCGCCCGGCAAGCA 2820
Qy 2692 GTGCTGTGATATCAATCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2751
Db 2821 GTGCTGTGATATCAATCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
Qy 2752 CTACACCTTGATCTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2811
Db 2881 CTACACCTTGATCTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
Qy 2812 AGGAAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2871
Db 2941 AGGAAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
Qy 2872 AAGAGATGATATTTTGTCTTTTGTGATTTGTGTATAGCCATCAATTCACAGAA 2931
Db 3001 AAGAGATGATATTTTGTCTTTTGTGATTTGTGTATAGCCATCAATTCACAGAA 3060
Qy 2932 ATTCCTTTCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2991
Db 3061 ATTCCTTTCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
Qy 2992 CAGGAGGATTAATTCACAGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 3051
```

```
Db 3121 CAGGAGGATTAATTCACAGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
Qy 3052 GTGTTTAACTCAGAGACCTTGTCTGTGTTTCCCTTAAAGATTTTGAAGCTGATAG 3111
Db 3181 GTGTTTAACTCAGAGACCTTGTCTGTGTTTCCCTTAAAGATTTTGAAGCTGATAG 3240
Qy 3112 TTGAGCATTTTATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 3171
Db 3241 TTGAGCATTTTATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 3300
Qy 3172 GCCAGCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3231
Db 3301 GCCAGCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
Qy 3232 AGCCAAAGACAAATATGAGCTCTTTTGAACAAACCTTGAATTTATTT 3283
Db 3361 AGCCAAAGACAAATATGAGCTCTTTTGAACAAACCTTGAATTTATTT 3412

RESULT 6
US-09-764-864-329
; Sequence 329, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT FILING DATE: US/09/764,864
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 329
; LENGTH: 3694
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-329

Query Match 93.0%; Score 3052.6; DB 9; Length 3694;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 3214; Conservative 1; Mismatches 1; Indels 132; Gaps 3;

Qy 68 GATATGATGAACAAAGCTTTATCATCGGAGAGCTGAGCCCGCGCTACCGCGAGAGAGCTTC 127
Db 91 GATATGATGAACAAAGCTTTATCATCGGAGAGCTGAGCCCGCGCTACCGCGAGAGAGCTTC 150
Qy 128 GAGAGCTCTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 187
Db 151 GAGAGCTCTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 210
Qy 188 AGGCTCTGTGAGACTACCCCGAGCAGAACTGGGCAATCGCGCATCGAGAGAGAGAGAG 247
Db 211 AGGCTCTGTGAGACTACCCCGAGCAGAACTGGGCAATCGCGCATCGAGAGAGAGAGAG 270
Qy 248 GTAAAGTGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 307
Db 271 GTAAAGTGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 330
Qy 308 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 367
Db 331 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 390
Qy 368 GACTTTTGTCTAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 427
Db 391 GACTTTTGTCTAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 450
Qy 428 CGGCGCTGTGACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 487
Db 451 CGGCGCTGTGACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510
Qy 488 TAAGCGGAGATCAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 547
Db 511 TAAGCGGAGATCAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 570
```

QY 548 TAGCTCCCTTCCTGCCCCCTCAGCAGAGCCCGTGGAGCCTTTCCCGGAGCAAG 607
Db 571 TAGCTCCCTTCCTGCCCCCTCAGCAGAGCCCGTGGAGCCTTTCCCGGAGCAAG 630
QY 608 GCCAGGCCCCCTGGGGGCACTTTCTCAGGCGCAGACAGATTGATTTCCCGCTGCGGATCTGG 667
Db 631 GCCAGGCCCCCTGGGGGCACTTTCTCAGGCGCAGACAGATTGATTTCCCGCTGCGGATCTGG 690
QY 668 TCCCCAGCCAGTTTGTGGTGCATCATCTGGAAGAAGAGGCTTTCACATTAAGAACATCA 727
Db 691 TCCCCAGCCAGTTTGTGGTGCATCATCTGGAAGAAGAGGCTTTCACATTAAGAACATCA 750
QY 728 CTAAACAGACCCAGTCCCGGAGTATATCATAGAAAAGAACTCTGAGAGCTGCAGAGA 787
Db 751 CTAAACAGACCCAGTCCCGGAGTATATCATAGAAAAGAACTCTGAGAGCTGCAGAGA 810
QY 788 AGCTGTACCATTCATGCGACCCCGAGGAGGACTTCTGAAGCATGCCGATGATTTCTTG 847
Db 811 AGCTGTACCATTCATGCGACCCCGAGGAGGACTTCTGAAGCATGCCGATGATTTCTTG 870
QY 848 AATCATGACAGAAAGAGGAGCAGATAGACCAAACTTCGCGAAGATTTCTCTGAAATCT 907
Db 871 AATCATGACAGAAAGAGGAGCAGATAGACCAAACTTCGCGAAGATTTCTCTGAAATCT 930
QY 908 TGGCACAATGAGCTTGGTGGAAAGCTGATTGGAAAAGAGGAGCAAAATTTGAAGAAA 967
Db 931 TGGCACAATGAGCTTGGTGGAAAGCTGATTGGAAAAGAGGAGCAAAATTTGAAGAAA 990
QY 968 TTGAACATGAACAGGAGCCAAAGTAAATCTCATCTTTGACAGATTGAGCATATACA 1027
Db 991 TTGAACATGAACAGGAGCCAAAGTAAATCTCATCTTTGACAGATTGAGCATATACA 1050
QY 1028 ACCGGAAGAGACATCATCTGTGAAGGCAAGTTGAGGCTGTGCGCATGATAG 1087
Db 1051 ACCGGAAGAGACATCATCTGTGAAGGCAAGTTGAGGCTGTGCGCATGATAG 1110
QY 1088 AGATTATGAAGAGAGCTGAGAGGCTTTGAAATGATATGCTGCTGTTAAC----- 1140
Db 1111 AGATTATGAAGAGAGCTGAGAGGCTTTGAAATGATATGCTGCTGTTAACCAAG 1170
QY 1141 ----- 1140
Db 1171 CCAATCTGATCCAGGCTTGAACCTCAGCGCATTTGGCATCTTTCAACAGAGACTGCG 1230
QY 1141 ----- 1140
Db 1231 TGCTATTCACACAGAGGCGCGGAGCTCCCGCTGCCCTACACCCCTTCA 1290
QY 1141 --ACCACTCCGATACTTCTCCAGCTGTACCCCATCAAGTTTGGCCCGTTCCGCG 1198
Db 1291 CTACCCGACTCGGATATTCTCCAGCTGTACCCCATCAAGTTTGGCCCGTTCCGCG 1350
QY 1199 ATCATCACTTTTCCAGAGCAGAGATGTGAATCTTTCAATCCAAACCCAGGCTGTG 1258
Db 1351 ATCATCACTTTTCCAGAGCAGAGATGTGAATCTTTCAATCCAAACCCAGGCTGTG 1410
QY 1259 GCGGCATCATCGGGAAGAGAGGAGCACAATCAAAAGCTGGGAGATTTGGCGAGCCT 1318
Db 1411 GCGGCATCATCGGGAAGAGAGGAGCACAATCAAAAGCTGGGAGATTTGGCGAGCCT 1470
QY 1319 CTATCAAGATTGCCCCCTGCGGAAGGCCAGACGTCAGCGAAAGATGTCATATCAACG 1378
Db 1471 CTATCAAGATTGCCCCCTGCGGAAGGCCAGACGTCAGCGAAAGATGTCATATCAACG 1530
QY 1379 GCGCACCGGAAGCCAGTTCAAGGCCAGGAGCGATCTTTGGGAAAATGAAAAGAGAAA 1438
Db 1531 GCGCACCGGAAGCCAGTTCAAGGCCAGGAGCGATCTTTGGGAAAATGAAAAGAGAAA 1590
QY 1439 ACTTCTTTAAACCCCAAGAGAGATGAAGTGAAGCGCATATCAAGATGCTCTTCCA 1498
Db 1591 ACTTCTTTAAACCCCAAGAGAGATGAAGTGAAGCGCATATCAAGATGCTCTTCCA 1650

QY 1499 CAGCTGGCCGGGTGATTGGCAAAAGGTGGCAAGACCGTGAACGACTGCAGAACTTAACCA 1558
Db 1651 CAGCTGGCCGGGTGATTGGCAAAAGGTGGCAAGACCGTGAACGACTGCAGAACTTAACCA 1710
QY 1559 GTGCAGAAATCATCTGCTCGTGACCAAAAGCCAGATGAAATAGGAAGATGCTCA 1618
Db 1711 GTGCAGAAATCATCTGCTCGTGACCAAAAGCCAGATGAAATAGGAAGATGCTCA 1770
QY 1619 GAATTATCGGCACTTCTTTGTAGCCAGACTGCAACGCGCAAGATCAGGAAATTTGAC 1678
Db 1771 GAATTATCGGCACTTCTTTGTAGCCAGACTGCAACGCGCAAGATCAGGAAATTTGAC 1830
QY 1679 AACAGTGAACAGAGAGAGAGAAATACCTCAGGAGGTGCTCTCAAGCGCAGCAAGT 1738
Db 1831 AACAGTGAACAGAGAGAGAGAAATACCTCAGGAGGTGCTCTCAAGCGCAGCAAGT 1890
QY 1739 GAGGCTCCACAGGACCCAGCAAAACAGATGAAATGATAGCCCTTCCAAACCTGACAG 1798
Db 1891 GAGGCTCCACAGGACCCAGCAAAACAGATGAAATGATAGCCCTTCCAAACCTGACAG 1950
QY 1799 AATGAGACCAACGACGACGACGATGGAGCAAAACCAAGACATCTGAGGAATGAG 1858
Db 1951 AATGAGACCAACGACGACGACGATGGAGCAAAACCAAGACATCTGAGGAATGAG 2010
QY 1859 AAGTCTGCGAGGCGGCGCAGGAGCTTGCAGAGCCCTGAGAACCCAGGAGCGAGAG 1918
Db 2011 AAGTCTGCGAGGCGGCGCAGGAGCTTGCAGAGCCCTGAGAACCCAGGAGCGAGAG 2070
QY 1919 GGGCGGGGAAGGTGAGCCAGGATTTGCCAGAACCCAGAGCCCGCTCCGCGCCCGAG 1978
Db 2071 GGGCGGGGAAGGTGAGCCAGGATTTGCCAGAACCCAGAGCCCGCTCCGCGCCCGAG 2130
QY 1979 GCTTCTGAGGCTTCAACCATCACTTCAACCATCACTCGATCTCTGAACTCCAC 2038
Db 2131 GCTTCTGAGGCTTCAACCATCACTTCAACCATCACTCGATCTCTGAACTCCAC 2190
QY 2039 GAGCTATCCCTTTAATTGAACTAATAGATGTAAGCTTTCAAGCCAAAGAAATGC 2098
Db 2191 GAGCTATCCCTTTAATTGAACTAATAGATGTAAGCTTTCAAGCCAAAGAAATGC 2250
QY 2099 ACAACCTTTTCTGTGCAATGCTCTGTACATGTTGTACATTTAGAAAGGAGAA 2158
Db 2251 ACAACCTTTTCTGTGCAATGCTCTGTACATGTTGTACATTTAGAAAGGAGAA 2310
QY 2159 TGTTAAGATATGAGCCTGTGAGTTACAGAGGTCCTGAGCGGTATATATTTAGAA 2218
Db 2311 TGTTAAGATATGAGCCTGTGAGTTACAGAGGTCCTGAGCGGTATATATTTAGAA 2370
QY 2219 ATATATATCAAAATCACTCACTCACTCAATTTTATCAATTAATTTTCT 2278
Db 2371 ATATATATCAAAATCACTCACTCACTCAATTTTATCAATTAATTTTCT 2430
QY 2279 TTTTAAAGAGAGAGGCTTTCTAGACTTTTAAAGAAATTAAGCTTTGGAGGCTCAC 2338
Db 2431 TTTTAAAGAGAGAGGCTTTCTAGACTTTTAAAGAAATTAAGCTTTGGAGGCTCAC 2490
QY 2339 GGTGTAGAGAGGCTTTGAGGCCACCCGCAAAATTCACCCAGAGGAAATCTGTCG 2398
Db 2491 GGTGTAGAGAGGCTTTGAGGCCACCCGCAAAATTCACCCAGAGGAAATCTGTCG 2550
QY 2399 GAGGACACTCAAGGAGATTTGATATCACTGTGTATGTCAAGAAAGGATACCGTCTC 2458
Db 2551 GAGGACACTCAAGGAGATTTGATATCACTGTGTATGTCAAGAAAGGATACCGTCTC 2610
QY 2459 CTTGAAGAGAACTCTGTACATCCCATGCGGTAGCTCATPACCCATTTCTCTT 2518
Db 2611 CTTGAAGAGAACTCTGTACATCCCATGCGGTAGCTCATPACCCATTTCTCTT 2670
QY 2519 GCTTCAAGGTTTAAACTGTTTGTGATATGCTATATTAATTTCTGTCTCTCTG 2578
Db 2671 GCTTCAAGGTTTAAACTGTTTGTGATATGCTATATTAATTTCTGTCTCTCTG 2730
QY 2579 TTTATCTCTCCCTCCCTCCCTCTCTCTCATCTCATCTTTTGAATTTCTC 2638


```

QY 2387 GAAATCTGTCGGAAGACACTCAGGCGAGTCTGGATCACTGTGATGTCACAGAG 2446
DB 1165 GAAATCTGTCGGAAGACACTCAGGCGAGTCTGGATCACTGTGATGTCACAGAG 1224
QY 2447 GATATCCGTCCTCTTGAAGAGAACTGTGATCACTGTCAGTCTGATGCTCAAC 2506
DB 1225 GATATCCGTCCTCTTGAAGAGAACTGTGATCACTGTCAGTCTGATGCTCAAC 1284
QY 2507 CCATTTCTTTGCTTCAAGGTTTAACTGGTTTTTGAATGCTATGCTATTAATTTCTC 2566
DB 1285 CCATTTCTTTGCTTCAAGGTTTAACTGGTTTTTGAATGCTATGCTATTAATTTCTC 1344
QY 2567 TGTCTCTCTGTTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2626
DB 1345 TGTCTCTCTG - TTAATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 1403
QY 2627 TTGAATTTCTGATCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 2684
DB 1404 TTGAATTTCTGATCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1463
QY 2685 CAAAGCAGTGTCTGAGTATCATCATCACAAAAGAGAAAGAGAGAGAGAGAGAGAGAG 2744
DB 1464 CAAAGCAGTGTCTGAGTATCATCATCACAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1523
QY 2745 GCCTCAACTTACACTTGTGTTACTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2804
DB 1524 GCCTCAACTTACACTTGTGTTACTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1583
QY 2805 TTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2864
DB 1584 TTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1643
QY 2865 CACATATGAAGATGATTTGTTGCTTTTGTGATTTTGTGATTTTGTGATTTTGTGAT 2924
DB 1644 CACATATGAAGATGATTTGTTGCTTTTGTGATTTTGTGATTTTGTGATTTTGTGAT 1703
QY 2925 CAAATATGATCTCTTCTTT - AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2983
DB 1704 CAAATATGATCTCTTCTTTTAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1763
QY 2984 TGTGGCCCGAGGCGTTAAATTCACAGATTTTAAAGAGAGAGAGAGAGAGAGAGAG 3043
DB 1764 TGTGGCCCGAGGCGTTAAATTCACAGATTTTAAAGAGAGAGAGAGAGAGAGAGAG 1823
QY 3044 TACCTCAGGTGTTTAACTGACAGCTGCTGTTGTTTCTGTTGATGATTTTGTAAA 3103
DB 1824 TACCTCAGGTGTTTAACTGACAGCTGCTGTTGTTTCTGTTGATGATTTTGTAAA 1883
QY 3104 GCTGATAGTTGAGACATTTTATTTTATTTTAAATGAATGAGTGAAGAGAGAGAG 3163
DB 1884 GCTGATAGTTGAGACATTTTATTTTATTTTAAATGAATGAGTGAAGAGAGAGAG 1943
QY 3164 TATCAACTGCGCAGCTGAGAGAGAGTGAAGTCAAGTGTGCAAGCTGTTCTGAATGT 3223
DB 1944 TATCAACTGCGCAGCTGAGAGAGAGTGAAGTCAAGTGTGCAAGCTGTTCTGAATGT 2003
QY 3224 CTTCGCTAGCGAAGAACATATGAGCTTTTGAACAACTTTGAAAATGTTTATTT 3283
DB 2004 CTTCGCTAGCGAAGAACATATATGAGCTTTTGAACAACTTTGAAAATGTTTATTT 2063

```

```

RESULT 8
US-10-262-445-39
; Sequence 39, Application US/10262445
; Publication No. US20040014058A1
; GENERAL INFORMATION:
; APPLICANT: Albrock II, John
; APPLICANT: Burgess, Catherine
; APPLICANT: Catterton, Elina
; APPLICANT: Chant, John
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Edinger, Shlomit

```

```

; APPLICANT: Gerlach, Valerie
; APPLICANT: Girot, Loic
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kexuda, Ramesh
; APPLICANT: Mezes, Peter
; APPLICANT: Millet, Isabelle
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Paturajan, Meera
; APPLICANT: Rieger, Daniel
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Zehrusen, Bryan
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Wei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE REFERENCE: 21402-462D
; CURRENT APPLICATION NUMBER: US/10/262,445
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/327,454
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/329,414
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/330,142
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/341,058
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/343,629
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: Cnaseqdist version 0.1
; SEQ ID NO 39
; LENGTH: 1707
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5)..(1669)
US-10-262-445-39

Query Match 47.6%; Score 1561.2; DB 16; Length 1707;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1624; Conservative 0; Mismatches 83; Indels 3; Gaps 1;

QY 66 ACGATGATGACAAAGCTTTTACATGCGGAACTGAGCCCGCGCTCAACCGCGGAGACT 125
DB 1 ACGATGATGACAAAGCTTTTACATGCGGAACTGAGCCCGCGCTCAACCGCGGAGACT 60
QY 126 CCGGAGCTCTTTGGGAGACGAGAGCTGCCCCGCGGGAGACAGGTCTGCTGAAGTCCGG 185
DB 61 CCGGAGCTCTTTGGGAGACGAGAGCTGCCCCGCGGGAGACAGGTCTGCTGAAGTCCGG 120
QY 186 CTACGCTTGTGAGACTACCCCGACAGAACTGGGCGCATCGGCGCATGAGAGACCCCTCTC 245
DB 121 CTACGCTTGTGAGACTACCCCGACAGAACTGGGCGCATCGGCGCATGAGAGACCCCTCTC 180
QY 246 GGGTAAAGTGAATTCATGAGAGAAATCATGAAAGTTGATTAATCAATGCTTAAAGCT 305
DB 181 GGGTCAAGTGAATTCATGAGAGAAATCATGAAAGTTGATTAATCAATGCTTAAAGCT 240
QY 306 AAGGACGAGAGAAATTCAGATTGGAACATCCCTCTCACTGCAATGAGAGGTGTTGA 365

```

Db 241 AAGAGCGAGAACATTCGATTCGAAATATCCCTCTCACTGACGTGGAGGTGTGA 300
QY 366 TGAATTTTGGCTCAATATGAGACAGTGAAGATGTGAAACAGTCAACACAGACAGA 425
Db 301 TGAATTTTGGCTCAATATGAGACAGTGAAGATGTGAAACAGTCAACACAGACAGA 360
QY 426 AACCCTGTTGCAACGTCACTATGCAACAGAAAGAAAGCAAAATATGATGATGAGA 485
Db 361 GACCCCTGTTGCAACGTCACTATGCAACAGAAAGAAAGCAAAATATGATGATGAGA 420
QY 486 GCTAAGCGGAGCATGATTTGAGAACTACTCTCTTCAAGATTTCTCAATCCCGATGAGA 545
Db 421 GCTAAGCGGAGCATGATTTGAGAACTACTCTCTTCAAGATTTCTCAATCCCGATGAGA 480
QY 546 GGTGAGCTCCCTTGGCCCTTCAAGCGAGCCAGCGTGGGAGCACTCTTCCCGGAGCA 605
Db 481 GGTGAGCTCCCTTGGCCCTTCAAGCGAGCCAGCGTGGGAGCACTCTTCCCGGAGCA 540
QY 606 AGGCCAGCGCCCTGGGGGAGCTTTCAGAGCGCAGAGATGATTTCCCGCTGGATCT 665
Db 541 AGGCCAGCGCCCTGGGGGAGCTTTCAGAGCGCAGAGATGATTTCCCGCTGGATCT 600
QY 666 GGTCCCAACCCAGTTTGTGTGTCATCATCGGAAAGAGGCGCTTGACCAATAAGAACAT 725
Db 601 GTTCCCAACCCAGTTTGTGTGTCATCATCGGAAAGAGGCGCTTGACCAATAAGAACAT 660
QY 726 CACTAAGCAGACCCAGTCCGSGATGATTCATATGAAAGAGAACTCTGAGCTGACAGA 785
Db 661 CACTAAGCAGACCCAGTCCGSGATGATTCATATGAAAGAGAACTCTGAGCTGACAGA 720
QY 786 GAAAGCTGTCACTATCCATGCAACCCAGAGGAGGAGCTTGTGAAGCATGCGGATGAT 845
Db 721 GAAAGCTGTCACTATCCATGCAACCCAGAGGAGGAGCTTGTGAAGCATGCGGATGAT 780
QY 846 TGAATCATGCAAGAAAGAGGAGATGAGCAAACTAGCGGAGAGATTTCTTGAAT 905
Db 781 TGAATCATGCAAGAAAGAGGAGATGAGCAAACTAGCGGAGAGATTTCTTGAAT 840
QY 906 CTGAGCACAATGCTGTGTGTGGAAGACTGATTTGAAAGAGAGGAGAAATTTGAAGA 965
Db 841 CTGAGCACAATGCTGTGTGTGGAAGACTGATTTGAAAGAGAGGAGAAATTTGAAGA 900
QY 966 AATTTGACATGAAAGAGGAGCCAAAGATAAATCTATCTTTCAGAGATTTGACATATA 1025
Db 901 AATTTGACATGAAAGAGGAGCCAAAGATAAATCTATCTTTCAGAGATTTGACATATA 960
QY 1026 CAACCCGAGAAAGACATCATCTGTGAAGGAGCAAGTTGAGGCTGTGCCAGTGTAGAT 1085
Db 961 CAACCCGAGAAAGACATCATCTGTGAAGGAGCAAGTTGAGGCTGTGCCAGTGTAGAT 1020
QY 1086 AGAGATTATGAGAGAGTGTGTGAGGCTTTGAAATATGATGTGTGTGTAACCCA 1145
Db 1021 AGAGATTATGAGAGAGTGTGTGAGGCTTTGAAATATGATGTGTGTGTAACCCA 1080
QY 1146 CTCCGATATCTTCTCAGCTGTATACCCCATCAACAGTTTGGCCGCTTCCCGATATCA 1205
Db 1081 CTCCGATATCTTCTCAGCTGTATACCCCATCAACAGTTTGGCCGCTTCCCGATATCA 1140
QY 1206 CTCTTATCCAGAGAGATGTGATCTCTTATCTCCAAACCCAGGCTGTGGGCGCAT 1265
Db 1141 CTCTTATCCAGAGAGATGTGATCTCTTATCTCCAAACCCAGGCTGTGGGCGCAT 1200
QY 1266 CATCGGAGAGAGGAGGAGCAATCAAAAGCTGTGGAGATTTGCCGAGAACCTTATCA 1325
Db 1201 CATCGGAGAGAGGAGGAGCAATCAAAAGCTGTGGAGATTTGCCGAGAACCTTATCA 1260
QY 1326 GATTTGCCCTGGAGAGGAGGAGCAAGCTCAAGAGAAAGATGATCTATCAACCGGAGCAC 1385
Db 1261 GATTTGCCCTGGAGAGGAGGAGCAAGCTCAAGAGAAAGATGATCTATCAACCGGAGCAC 1317
QY 1386 GGAAGCCAGTTCAGAGGAGGAGGAGCAAGATCTTTGGGAAACTGAAAGAGGAAATCTTCTT 1445

Db 1318 GGAATCCCACTTCAAGGCCCAAGGACGAGATCTTTGGAAACTGAAAGAGAAATCTTTT 1377
QY 1446 TAAOCCCAAG 1505
Db 1378 TAAOCCCAAG 1437
QY 1506 CCGGAGATTTGGCAAAAGGTGGCAAGCCGTGAACGAATCTGCAAACTTAAACAGTGCAGA 1565
Db 1438 CCGGAGATTTGGCAAAAGGTGGCAAGCCGTGAACGAATCTGCAAACTTAAACAGTGCAGA 1497
QY 1566 AGTCACTGTGCTCGTGCACAAAGCCAGATGAAATATGAGAAAGATGATGATGATAT 1625
Db 1498 AGTCACTGTGCTCGTGCACAAAGCCAGATGAAATATGAGAAATGATGATGATGATAT 1557
QY 1626 CCGGAGATTTCTTTGTATGACAGAGTGCACAGGAGAGATGAGGAAATTTGACACAGGT 1685
Db 1558 CCGGAGATTTCTTTGTATGACAGAGTGCACAGGAGAGATGAGGAAATTTGACACAGGT 1617
QY 1686 GAAAGCAG 1745
Db 1618 GAAAGCAG 1677
QY 1746 CCACAGGACACGCAAAACCAACGATGAT 1775
Db 1678 CCACAGGACACGCAAAACCAACGATGAT 1707

RESULT 9
US-09-764-864-749
; Sequence 749, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 749
; LENGTH: 1186
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (511)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (774)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (777)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-749

Query Match 33.4%; Score 1096.8; DB 9; Length 1186;
Best Local Similarity 98.4%; Pred. No. 1.7e-289;
Matches 1145; Conservative 3; Mismatches 12; Indels 4; Gaps 4;

QY 1247 CCCAGGCTGTGGGCGGCATATGAGGAGAGAGGAGGAGAGAGAGAGAGAGAGAGAT 1306
Db 25 CCCAGGCTGTGGGCGGCATATGAGGAGAGAGGAGGAGAGAGAGAGAGAGAGAGAT 84
QY 1307 TCGCGGAGGCTCTTATCAAGATTTGCCCTCGGAGGAGGAGAGAGAGAGAGATG 1366
Db 85 TCGCGGAGGCTCTTATCAAGATTTGCCCTCGGAGGAGGAGAGAGAGAGAGATG 144
QY 1367 TCATCATCACCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCTTTGGAAAC 1426
Db 145 TCATCATCACCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCTTTGGAAAC 203
QY 1427 TGAAGAGAGAAATCTTTTAAACCCCAAGAGAGAGAGAGAGAGAGAGAGATCAAG 1486

Db 204 TGAAGAGGAAAACTCTTTAAACCCCAAGAGAGTGAAGCCATTATCAGAG 263
Qy 1487 TGCCCTCTTCCACAGCTGGCGGGGATTTGGAAAGTGGCAAGACCGTGAACGAACTGC 1546
Db 264 TGCCCTCTTCCACAGCTGGCGGGGATTTGGAAAGTGGCAAGACCGTGAACGAACTGC 323
Qy 1547 AGAATTAAACAGTGCAGAGTGCATGCTGCTCTGTGACCAAAAGCCAGTGAATGAG 1606
Db 324 AGAATTAAACAGTGCAGAGTGCATGCTGCTCTGTGACCAAAAGCCAGTGAATGAG 383
Qy 1607 AAGTATGTCAGAAATTAATCGGGCACTTCTTTGCTAGCCAGACTGCAAGCGCAAGTCA 1666
Db 384 AAGTATGTCAGAAATTAATCGGGCACTTCTTTGCTAGCCAGACTGCAAGCGCAAGTCA 443
Qy 1667 GGGAAATTTG-TACAAAGTGTAGAGAGAGAGAGCAAAATCCTCAAGGAGTGCCTCA 1725
Db 444 GGGAAATTTGTACAAAGTGTAGAGAGAGAGAGCAAAATCCTCAAGGAGTGCCTCA 503
Qy 1726 CAGGCGACAGAGTGGCTCCCAAGGCGACAGCAAAACAAAGGATGTAAGCCCTTC 1785
Db 504 CAGGCGA-NAAGTGAAGCTCCCAAGGCGACAGCAAAACAAAGGATGTAAGCCCTTC 562
Qy 1786 CAACACTGACAGATGAGACCAACGACGACAGCCAGATCGGAGCAAAACCAAGCA 1845
Db 563 CAACACTGACAGATGAGACCAACGACGACAGCCAGATCGGAGCAAAACCAAGCA 622
Qy 1846 TCTGAGGAATGAGAGTCTGCGAGGCGCGCAGGAGCTCTCCAGAGGCTCTGAGAACCC 1905
Db 623 TCTGAGGAATGAGAGTCTGCGAGGCGCGCAGGAGCTCTCCAGAGGCTCTGAGAACCC 682
Qy 1906 AGGGGCGAGAGGGGCGGGAGAGTCAAGCAGGTTTGCGAGAACCAACGAGCCCGCT 1965
Db 683 AGGGGCGAGAGGGGCGGGAGAGTCAAGCAGGTTTGCGAGAACCAACGAGCCCGCT 742
Qy 1966 CCGGCCCCCAAGGCTCTGCGAGGCTTCAAGCAT-CACTTCAACATCACTCGGATCTC 2024
Db 743 CCGGCCCCCAAGGCTCTGCGAGGCTTCAAGCATCACTTCAACATCGGATCTC 802
Qy 2025 TCCGAACTCCCAAGAGCTTATCCCTTTTATGTTAACTAATATGTTGAACTGTTCAA 2084
Db 803 TCCGAACTCCCAAGAGCTTATCCCTTTTATGTTAACTAATATGTTGAACTGTTCAA 862
Qy 2085 GCCAAGCAATGACACCCCTTTTCTGTGCAAAATGCTCTGTATACATGTTGTAATA 2144
Db 863 GCCAAGCAATGACACCCCTTTTCTGTGCAAAATGCTCTGTATACATGTTGTAATA 922
Qy 2145 TTGAAAGGGAAGATGTTAAGATATGAGCTGTGAGTTACAGAGGTCCTGAGCGT 2204
Db 923 TTGAAAGGGAAGATGTTAAGATATGAGCTGTGAGTTACAGAGGTCCTGAGCGT 982
Qy 2205 AATATATTTTGAATATATATCAATACTCACTAATCTCAATTTTAACTAATAT 2264
Db 983 AATATATTTTGAATATATATCAATACTCACTAATCTCAATTTTAACTAATAT 1042
Qy 2265 TAAATTTTCTTTTAAAGGAAAGAGGCTTTCTAGACTTTTAAATTAATTAAT 2324
Db 1043 TAAATTTTCTTTTAAAGGAAAGAGGCTTTCTAGACTTTTAAATTAATTAAT 1102
Qy 2325 TTGGAGGTCTCAACGCTGTAGAGAGAGCTTTGAGGCAACCGCACAATAATCAACCA 2384
Db 1103 TTGGAGGTCTCAACGCTGTAGAGAGAGCTTTGAGGCAACCGCACAATAATCAACCA 1162
Qy 2385 GGGAAATCTGTCGGAAAGCACT 2408
Db 1163 GGGAAATCTGTCGGAAAGCACT 1186

RESULT 10
US-09-735-705-347
; Sequence 347, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
APPLICANT: Pan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aljun
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
APPLICANT: Fanger, Neil
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: AND DIAGNOSIS OF LUNG CANCER
CURRENT FILING DATE: 21021.455C14
CURRENT FILING DATE: US/09/735,705
NUMBER OF SEQ ID NOS: 2000-12-12
SOFTWARE: PatsSeq for Windows Version 3.0
SEQ ID NO 347
LENGTH: 1740
TYPE: DNA
ORGANISM: Homo sapiens
US-09-735-705-347
Query Match 21.2%; Score 697.2; DB 9; Length 1740;
Best Local Similarity 64.9%; Pred. No. 1e-179;
Matches 1116; Conservative 0; Mismatches 538; Indels 66; Gaps 3;
Qy 73 ATGAAACAAGCTTTTATCATCGGGAACCTGAGCCCGCTGACCGCGAGACTCCGAG 132
Db 1 ATGAAACAAGCTTTTATCATCGGGAACCTGAGCGAGAACCGCGCTTGGACTTGAAGT 60
Qy 133 CTCTTTGGGAGCAAGAGAGTGCCTGCGGAGAGCTGCTGCTGAGATCCCGCTACGCC 192
Db 61 ATCTTAAAGACCGCAAGATCCCGGTTCGGAGACCTTCTGTGAGAGACTGCTACGCC 120
Qy 193 TTGCTGACTACCCCGACAGAACTGGGCCATCCGCGCATCGAAGCCCTCTCGGTAA 252
Db 121 TTGCTGACTGCGCGGACAGAGAGCTGGCCCTCAAGGCATCGAGGCGCTTCAAGTAA 180
Qy 253 GTGGAATTGATGAGGAAATCATGAGAGTGTATCTCATGCTTAAAGCTTAAGAGC 312
Db 181 ATGAACTGACGAGAAACCCATGAGAGTGTAGACATCGGTCCCAAAAGCGAAAGATT 240
Qy 313 AGGAAATTCAGATTGAGAAATCATCCCTCTCACTGAGTGGAGAGTGTGATGACTT 372
Db 241 CGGAACTTCAGATTGAGAAATCATCCCGCTCATTTACAGTGGAGAGTGTGATGTTA 300
Qy 373 TTGCTCATATGAGACAGTGAAGATGTGAAACAACTCAACACAGCAAGAAACGCC 432
Db 301 CTAGTCAGATGAGAGTGTGAGAGCTGTGACCAAGTGAACCTGACCGAAACTGCA 360
Qy 433 GTTGTCAAGTCAATATGAGCAACAGAGAAAGCAAAATAGCATGAGAGACTTAAC 492
Db 361 GTTGTCAAGTCAATATGAGCAATATGAGCAACAGAGAAAGCAAAATAGCATGAG 420
Qy 493 GGGCATCAGTTTGAAGACTACTCTCAAGATTCTTCACTCCCGATGAGAGGTAGC 552
Db 421 GGAATTCAGTTTGAAGATTCTTCAAGTTGAGAGTGTGATCTTCTGATGAGAGGCGCC 480
Qy 553 TCCCTTCCGCTTCAAGGAGCCCAAGCTGGGAGCACTCTTCCCGGAGCAAGGC--- 609
Db 481 CAGCAAAACCCCTTTCAGAGAGCCCGGAGGTGCGCGGGGCTTGGGAGAGGGGCTCTCA 540
Qy 610 -----CAGGCGCTGGGGGCACTTCTCAGGCGCAAGCATTAATTTCCGCTGGATC 663
Db 541 AGCAGGGGTCTTCAGAGATCCGATCAAGCAAGAAACATGTGATTTGCTTGGCGCTG 600
Qy 664 CTGGTCCCAACCAAGTTTGTGAGTGCATCATCGAAAGAGAGGCTTGAACATTAAGAAC 723
Db 601 CTGGTCCCAACCAATTTTGTGAGTGCATCATGAGAAAGAGGTGCAACATTCGGAAC 660


```

Db      661 ATCAACAAGACCCAGTCTAAATGATGTCCACCGTAAAGAAATGGCGGGCTGCT 720
QY      784 GAGAGGCTGTGACATTCATGATCCACCCAGAGGGGACTTGTGAAGATGCCGATGTT 843
Db      721 GAGAGGCTGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
QY      844 CTGGAATTCATGCGAAGAGGAGGAGATGAGCAAACTAGCGAAGAGATCTCTGAAA 903
Db      781 CTGGAATTCATGCGAAGAGGAGGAGATGAGCAAACTAGCGAAGAGATCTCTGAAA 840
QY      904 ATCTTGACACAAATGCTGTGTGTGAAGACTGATTTGAAAAAGAGGCAAAATTTGAAG 963
Db      841 ATTTAGTCAATTAATTAATTTGTTGGAAGTCTTAATTTGTTAAAGAAAGAAATCTTAAA 900
QY      964 AAAATTGAACTGAAACAGGAGCAAGTAACATCTATCTTTGACAGATTTTGACATA 1023
Db      901 AAAATTGAACTGAAACAGGAGCAAGTAACATCTATCTTTGACAGATTTTGACATA 960
QY      1024 TACAAACCGGAAAGACATCATCTGTGAAGGGCAAGTTGAGGCTGTGAGCTGAG 1083
Db      961 TATAATCCAGAACGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
QY      1084 ATAGAGATTATGAAGAACTGCTGTGAGGCTTTGAAAAATGATATGCTGTGTTAACAC 1143
Db      1021 GAGAGGATCATGAAGAAATCAAGGAGCTTATGAAGAAATGATATGCTGTGTTAACAT 1080
QY      1144 CACTCCGAT-----ACTTCTCA 1162
Db      1081 CAAGCACATTTAATTCGTGATTTAATCTGAACGCTTGGGTCTGTCTCCACCACTTCA 1140
QY      1163 GCTTACCCCCATGACCAAGTTTGGCCGTTCCCGCATCATCTC----- 1208
Db      1141 GGGATGCCACTCCCACTCAGAGGCCCTTACGACATGACCTCTCCACCGCAGTTT 1200
QY      1209 --TTATCCAGACAGAGATTTGATCTCTTATCCCAACCGCTGTGAGGCGCATC 1266
Db      1201 GAGCAATCAAGAAAGAGAGCTGTTCATCTGTTATCCGAGCTTATCACTGCTGCTCAT 1260
QY      1267 ATCGGAAAGAGGGGACACATCAACAGCTGGGAGATTGGCCGAGGCTTATCAAG 1326
Db      1261 ATCGGAAAGAGGGGACACATCAACAGCTGGGAGATTGGCCGAGGCTTATCAAG 1320
QY      1327 ATTTGCCCCGAGAGGAGGAGGAGCTCAGCGAAAGGATGATCATCAACCGGCGACCG 1386
Db      1321 ATTTGCTCCAGCGAGACACATGCTAAAGAGATGATGATTAATCACTGGAACACCA 1380
QY      1387 GAAGCCCACTTCAAGGCCAGGAGCGATCTTTGGGAAACTGAAGAGGAAATCTTTT 1446
Db      1381 GAGGCTCACTTCAAGGCTCAGGGAAGAAATTTATGAAAAATTAAGAAAAAATCTTTGTT 1440
QY      1447 AACCCCAAGAAAGTGAAGCTGGAAGGCAATTCAGAGTCCCTTCCACAGCTGGC 1506
Db      1441 ACTCCCTAAAGAGGATGAACTTGAAGCTCATATAGAGTGCATCTTTGTCTGAGC 1500
QY      1507 CGGGGATTTGGCAAGAGTGGCAAGACGTAAGCAAGTCAAGAACTTAAACAGTGCAGAA 1566
Db      1501 AGAGTTATTGAAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY      1567 GTCATGCTGCTGTGACCAAGCGCAGATGAAGAAATGAAGAAATGATCTGAGAAATATC 1626
Db      1561 GTTGTGTCTGCTGTGACCAAGCACTGATGAAGAAATGAAGAAATGATCTGAGAAAT 1620
QY      1627 GGGCACTTTCTTGTGCTGAGCAAGTGAAGCGGCAAGTCAAGGAAATTTGAACAAAGG 1686
Db      1621 GGTCACTTTCTTGTGCTGAGCAAGTGAAGCGGCAAGTCAAGGAAATTTGAACAAAG 1680
QY      1687 AAGCAGAGGAGCAAAATACCTCTCAGGAGTGCCTCAG 1726
Db      1681 AAGCAGAGGAGCAAAATACCTCTCAGGAGTGCCTCAG 1720

```

```

; Sequence 347, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margareta
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.45C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-007-700-347

```

```

Query Match      21.2% Score 697.2; DB 13; Length 1740;
Best Local Similarity 64.9%; Pred. No. 1e-179;
Matches 1116; Conservative 0; Mismatches 538; Indels 66; Gaps 3;

QY      73 ATGAACAAGCTTTACATCGGAACTGAGCCCGCGTCAACCGGAGCACTCCGGCAG 132
Db      1 ATGAACAAGCTTTATATGGAACCTCAGCGAAGCGCCCGCTCGGACCTAGAAAGT 60
QY      133 CTCTTTGGGAGAGAGAGCTGCCCCCTGGCGGAGCAGTCTGCTGAAGTCCGGCTAAGCC 192
Db      61 ATCTTCAGAGAGCGCAAGATCCGGGTGAGACCTTCTGTGAGACTGAGCTACGCG 120
QY      193 TTGCTGACTACCCGACAGAACTGGGCTATCCGGCCATCGGCGCATGAGACCTCTCGGTAAA 252
Db      121 TTGCTGACTGCGCGAGAGAGCTGGGCTCTGAGGCATGAGGCGCTTCAAGGTAAA 180
QY      253 GTGGAATTGCAATGGGAAATCAATGAAATGATTAATCAATCTTAATAAAGTAAAGAGC 312
Db      181 ATGAAGCTCAGCGGAAACCCATAGAAATGAGCACTCGATCCCAAAAAGGCAAGGAT 240
QY      313 AGGAAATTCAGATTTGAAACATCCCTCTCTCACTGCACTGGGAGGCTGTGATGACTT 372
Db      241 CGGAAATTCAGATTTGAAATATCCGCTCTCAATTAATGAGGAGGCTGTGATGACTT 300
QY      373 TTGCTCAATTTAGGACAGTGAAGATGTGAACAAGTCAACAGACACAGAAACCGCC 432
Db      301 CTAGTCAATTAAGAGTGTGAGAGCTGTGAGCAAGTGAACATCACTCGGAAATGCA 360
QY      433 GTTGTCAACGTCAATATGCAACAGAGAAAGAAATTAAGCCATGTGAGAAAGCTAAGC 492
Db      361 GTTGTAAATGTAACCTATTCAGTAAAGGACCAAGCTAGCAAGCACTAAGCAATGAAT 420
QY      493 GGGCATCAGTTTGAAGAACTACCTTCAAGATTTCTCAATCCCGGATGAAGAGTGAAGC 552
Db      421 GGAATTCAGTTAAGAAATTCACCTTGAAGATGCTTAATTCCTTAATGAAGAGCGCGCC 480
QY      553 TCCCTTGGCCCGCTCAGCGAGCCAGCGTGGGAGCACTTTCGCGGAGCAAGGCG-- 609
Db      481 CAGCAAAACCCCTTGAAGAGCGCCGAGAGTCCGCGGGGCTTGGGAGAGAGGCGCTCTCA 540

```

QY 610 -----CAGGCCCTGGGGGCACTTCTCAGGCCAGACAGATGATTTCCCGCTGGGATC 663
Db 541 AGGCAAGGGGTCTCCAGGATCCGTATCCAGAGCAAAACATGATGATTTCCCTCTGGGCTG 600
QY 664 CTGGTCCCCACCCAGTTGTTGGTCCATCATCGGAAAGAGGGCTTGACATTAAGAC 723
Db 601 CTGGTCCCCACCCAGTTGTTGGTCCATCATCGGAAAGAGGGCTTGACATTAAGAC 660
QY 724 ATCACTAGCAGACCCAGTCCCGGTGATATCCATTAAGAAAGAACTCGAGCTGCA 783
Db 661 ATCACTAGCAGACCCAGTCCCGGTGATATCCATTAAGAAAGAACTCGAGCTGCA 720
QY 784 GAGAGCCTGTACCATCATGCAATCCAGAGGGGAACTTGAAGATGCGCATGAT 843
Db 721 GAGAGCCTGTACCATCATGCAATCCAGAGGGGAACTTGAAGATGCGCATGAT 780
QY 844 CTGGAATCATGCAAGAAAGAGGCAAGTGAACAACTAGCCGAAGATTTCTTGAAA 903
Db 781 CTGGAAGATTAAGCAATTAAGAGGCTCAAGATTAATAATTCAGAGAGATCCCTTGAAG 840
QY 904 ATCTTGACACCAATGGCTTGGTGAAGATGATGGAAGAAAGAGCAAAATTTGAG 963
Db 841 ATTTAGCTATTAATTAATTTGTTGGAGCTTATGTTAAAGAGAAATCTTAAA 900
QY 964 AAAATTGACATGAAAGAGGACCAAGTAAACAATCTCATCTTTCAGATTTGACATA 1023
Db 901 AAAATTGACATGAAAGAGGACCAAGTAAACAATCTCATCTTTCAGATTTGACATA 960
QY 1024 TACACCCCGGAAAGACATCATCTGTGAAGGGCAAGTTGAGGCTGTGCAAGTGTAG 1083
Db 961 TATATCCAGAACCCACTATTACAGTTAAAGGCAATGTGAGCATGTGCCAAAGCTGAG 1020
QY 1084 ATAGATTAATGAAGAGCTGCGGCTTGAAGGCTTGAAGATGATGCTGGCTGTTAAAC 1143
Db 1021 GAGAGATTCATGAAGAAATCAAGGAGCTTATGAATAATGATGCTTATATAACTT 1080
QY 1144 CACTCCGAT-----ACTTCTCA 1162
Db 1081 CAGGCAATTAATTCCTGATTAATCTGAAGCCTTGGGTCTGTTCCCAACCATCTTA 1140
QY 1163 GCTGTACCCCATCATCAAGTTGGCCGCTTCCGATCATCTC-----1208
Db 1141 GGGATGACACTCCCATCTCAGGGCCCTTTCAGCCATGACTCTCTCCCTAACCGCAGTT 1200
QY 1209 --TTATCAGAGCAGGATTTGAATCTCTCATCCCAACCGAGGTGTGGGCGCATC 1266
Db 1201 GAGCATCAAGAAACGAGAGCTGTCTATCTGTTATCCAGCTTATCAAGTGGGCTATC 1260
QY 1267 ATCGGAAAGAGGGGACATCAATCAACAGCTGGCAGATTTGCGCGAGCCTCATCAAG 1326
Db 1261 ATCGGAAAGAGGGGACATCAATCAACAGCTGGCAGATTTGCGCGAGCCTCATCAAG 1320
QY 1327 ATTGCCCCCTCGGAGGCGCAGAGCTCAGGAAAGATGTCATCACCGGCGCACG 1386
Db 1321 ATTGCTCCACCGGAGAGCAGAGTGTAAAGTGAGATGATTAATCACTGAGCACCA 1380
QY 1387 GAAGCCAGTTCAAGGCGCCAGGAGCGATCTTTGGGAACTGAAAGGAAATCTTT 1446
Db 1381 GAAGCTCAAGTTCAAGGCTCAGGAAAGATTTATGAAATTAAGAAAGAAATCTTTGTT 1440
QY 1447 AACCCAAAGAAAGTGAAGTGAAGCGCATATCAGATGCGCTTTTCCAGAGCTGCG 1506
Db 1441 AGTCTAAAGAAAGAGTGAAGTGAAGCTCATATCAGATGCGCTTTTCTGCTGCGC 1500
QY 1507 CGGGTATTGGCAAGGTGGCAAGCCGTGAAGCACTTGAGAACTTTAACAGTGCAGAA 1566
Db 1501 AGAGTTATTGGAAGAGGCGCAAAACGCTGAATGAATCTCAGAAATTTGCAAGTGCAGAA 1560
QY 1567 GTATGCTGCTGTGACCAAGCGCATGAAATTAAGAAAGTATGCTGCAAAATATTC 1626
Db 1561 GTTGTGTCTCTGTGACCAAGCGCTGATGAATATCAAGAGTGTGTCAAATAACT 1620
QY 1627 GGGCACTTTCTTCTGACCAAGCTGCAAGCGCAAGATCAGGAAATTTGTAACAGGTG 1686

Db 1621 GGTACTTCTATGCTTGGCAGGTTGCCAGGAAATTCGAAATTTCTGACTCAGSTA 1680
QY 1687 AAGCAGAGAGCAGAAATACCTCAGGAGTCCGCTCAC 1726
Db 1681 AAGCAGAGCAGCAACAGAAAGCTCTGCAAGTGCACAC 1720

RESULT 14
US-10-117-982-347
; Sequence 347, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kaios, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-982-347

Query Match 21.2%; Score 697.2; DB 15; Length 1740;
Best Local Similarity 64.9%; Pred. No. 1e-179;
Matches 116; Conservative 0; Mismatches 538; Indels 66; Gaps 3;

QY 73 ATGAACAAGCTTTATCATCGGAAACCTGAGCCCGCTCACCGCGAGACCTCCGCGAG 132
Db 1 ATGAACAAGCTTTATCGGAAACCTGAGAGAGCGCGCCCTCGGACTGAGAAAGT 60
QY 133 CTCTTGGGAGCAGGAAGCTGCCCCCTGGCGGACAGTCTCTGTGAAGTCCGAGTAC 192
Db 61 ATCTTCAAGGACCGCAAGATCCCGGTGTGGGACCTTCTGTGGAAGATGAGTACGCG 120
QY 193 TTGTTGACATACCCGACCAAGAACTGGGCGATCCGCGCATCGAGACCTCTCGGTTAA 252
Db 121 TTGTTGACATGCCCCGAGCGAGAGCTGGCCCTCAAGGCCATCGAGGCGCTTTCAGTTAA 180
QY 253 GTGAATTTGATGGGAAATCATGAAATGATTAATCATCTCTTAAAGGCTAAGAGC 312
Db 181 ATGAAGTGCAGCGGAAACCATAGAAATTTGAGACTCGGTTCCAAAAGGCAAGAT 240
QY 313 AGGAAATTCAGATTGAGAAATCATCCCTCTCACTGAGTGGAGGATGATGAGACTT 372
Db 241 CGGAAATTCAGATTGAGAAATCATCCCGCTCATTTACAGTGGAGGATGATGATTTA 300
QY 373 TTGGCTCATATGAGACATGAGAAATGTTGAAACAACTCAACAGACAGAAACCGCC 432
Db 301 CTAGTCCAGTATGAGATGTTGAGAGCTGTGACCAAGTGAACCTGACGAAATCGCA 360
QY 433 GTTGTCAAGCTCATATGCAACAGAGAAAGGAAATTAAGCAATGAGAAAGTAAAG 492
Db 361 GTTGTAAATGTAATCTATTCATTAAGACCAAGCTGAGAAAGCACTGACAAATTAAT 420
QY 493 GGGCATCAAGTTTGAAGACTATCTCTTCAAGATTTCTTCAATCCCGATGAAGAGTGAAGC 552
Db 421 GGAATTCAGTTGAGAAATTCATCTGAAAGTATGCTATATCTCTGATGAAAGGCGCGCC 480

```

QY 553 TCCTCTGCCCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGCAAGGC--- 609
Db 481 CAGCAAAACCCCTTGAGAGAGCCGAGGTGCGGGGGCTTGGGCAAGGGGCTCTCA 540
QY 610 -----CAGCGCCCTGGGGGCACTTCTCAGGCGCAGACAGATTGATTTTCCGCTGGATC 663
Db 541 AGGCAAGGGGTCTCAGAGATCCGTATCCAGCAAAACAGATGATTTGCTCTGGGCTG 600
QY 664 CTGGTCCCAACCCAGTTTGTGTGCTCATTCGAAAGAGGGCTTGAACATAAAGAC 723
Db 601 CTGGTCCCAACCCAGTTTGTGTGCTCATTCGAAAGAGGGCTTGAACATAAAGAC 660
QY 724 ATCACTAGAGAGACCCAGTCCCGGGTAGATATCCATAGAAAAGAACTCTGAGCTCA 783
Db 661 ATCACTAGAGAGACCCAGTCTTAATTCATGTCCACCTTAAGAAAATGCGGGGCTCT 720
QY 784 GAGAGGCTGTACCCATCCATGCCACCCAGAGGGGACTTTCGAAAGATGCCGATGAT 843
Db 721 GAGAGGCTGTACATCTCTCTATCTCTGAAAGGCACTCTGGCGCTTGAAGCTAT 780
QY 844 CTGGAATCATGCAAGAGGAGATGAGACCAACTAGCCGAGAGATTCCTCTGAAA 903
Db 781 CTGGAGATTATGCAAGAGGCTCAGATATTAATTCACAGAGAGATCCCTTGAAG 840
QY 904 ATCTTGGACAACAATGGCTTGGTGAAGACTGATTGGAAAAGAGCAAAATTGAAG 963
Db 841 ATTTAGCTCATTAATCTTTTGTGAGCTCTTAATTTGGTAAGAAAGAAATCTTAAA 900
QY 964 AAAATTTGAACATGAACAGAGCAACAATTAACAATCTCATCTTTCAGAGATTGACATA 1023
Db 901 AAAATTTGAACAGCAACAGACCTAAATCAGATATCTCCATTCAGAGAAATTGACGCTG 960
QY 1024 TACAAACCCGGAAGAACCTCATCTGTGAAGGCAAGTTCGAGCCGAGCTGAG 1083
Db 961 TATAATCCAGAACGACATTAACAAGTAAAGGCAATGTGAACAATGTCMAAGCTGAG 1020
QY 1084 ATAGAGATTATGAGAGAGTGGGTGAGGCTTTCGAAATGATATGCTGTTAAACAC 1143
Db 1021 GAGAGATATATGAAGAAATCAGGAGTCTTAAGAAATGATATGCTTATGAAATCTT 1080
QY 1144 CACTCCGAT-----ACTTCTCCA 1162
Db 1081 CAGACATATTAATCTGATTAATCTGAACGCTTGGTCTGTTCCCACTTCA 1140
QY 1163 GCTGTACCCCATCACCAGTTTGGCCGCTTCCGATCATCACTC----- 1208
Db 1141 GGGATGCCACTTCCACTCAGGGGCCCTTCAAGCCATATCTCTCCTAACCCGAGTTT 1200
QY 1209 --TTATCCAGAGAGATTTGAAATCTTTCATCTTCAACCAAGGCTGTGGGCGCATC 1266
Db 1201 GAGCAATCGAAGAGAGACTGTTCATCTGTTTATCCAGCTCTCATAGTGGTGCATC 1260
QY 1267 ATCGGAGAGAGGGGACACATCAAAAGCTGGCGAGATTGGCGGAGCTCTATCANG 1326
Db 1261 ATCGGAGAGAGGGGACACATCAAGCAAGCTTCTCGCTTGTGAGACTTCAATTAAG 1320
QY 1327 ATTCGCCCCCGGAGAGCCCAAGCTCAGCGAAAGGATGCTATCATCAACCGGGCCACCG 1386
Db 1321 ATTCGCCCCCGGAGAGCCCAAGTGTCTAAAGTGAAGTGTGATTAATCATCTGACCA 1380
QY 1387 GAGAGCCAGTTCAAGGCCAGGAGCGGATCTTTGGGAACTGAAAGAGGAAATCTTCTT 1446
Db 1381 GAGGCTCAGTTCAAGGCTCAGGGAAGATTTATGAAAAATTTAAAGAAAGAAATCTTGT 1440
QY 1447 AACCCCAAGAGAGAGTGAAGCTGAAGCGCATATCAAGTGCCTCTTCCAGAGCTGGC 1506
Db 1441 AGTCCCTAAGAGAGAGTGAAGCTTGAAGCTCATATCAGAGTGCATCTTGTGCTGGC 1500
QY 1507 CGGCTATTTGCAAGAGTGGCAAGACCGTGAAGAACTGCAAGAACTTAACCAATGAGAA 1566
Db 1501 AAGATTAATTTGAAGAGAGCAAGAAAGGTAAGAACTTCAAGAAATTTGTAATGAGAA 1560
QY 1567 GTCATGTGCTGTGACCAACGCGAGATGAAGAAATGAGAAAGTGAATCGTCAGAAATATC 1626

```

```

Db 1561 GTTGTGCTCTCGTACAGACACTGATGAGATGACCAAGTGTGTGTAATAATTAATCT 1620
QY 1627 GGGCACTTCTTTGTGTAAGCAATGCAACAGGCAAGATCAGGAAATGTCAACAGGTG 1686
Db 1621 GGTCACTTCTATGTCTTGGCAGGTGGCCAGAAAAAATTCAGGAAATTCGATCTCAGGTA 1680
QY 1687 AAGCAGCAGAGAGAAATATCCCTCAGGAGTGCCTCAC 1726
Db 1681 AAGCAGCAGCAACAACAGAAAGCTCTGCAAAAGTGAACAC 1720

RESULT 15
US-10-117-982-478
; Sequence 478, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Maranabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kallos, Michael D.
; APPLICANT: Meticle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 478
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-117-982-478

Query Match 21.2%; Score 697.2; DB 15; Length 1740;
Best Local Similarity 64.9%; Pred. No. 1e-179; Indels 66; Gaps 3;
Matches 1116; Conservative 0; Mismatches 538;

```

QY 493 GGGCATGATTGAGAACTACTCTTCAAGATTTCTACATCCGGATGAAAGGTGAGC 552
DB 421 GGATTTGAGTTAGAAATTTACCTTTAAAGTAGCTTATATCCGTAGTAAGACGGCCCC 480
QY 553 TCCCTTGGCCCCCTCAGCGAGCCAGGCTGAGGACCACTCTTCCCGGAGCAAGGC--- 609
DB 481 CAGCAAAACCCCTTGACACAGCCGAGGTGCGCGGGGCTTGGGCAAGAGGGCTCTCA 540
QY 610 -----CAGCGCCCTGCGGGGCACTTCTCAGGCGCAGACGATTTGATTTCCCTGCGATC 663
DB 541 AGGAGAGGGGTCTCCAGGATCCGTATCCAGAGAAACCATGTGATTTGCCCTGCGCTG 600
QY 664 CTGCTCCCAACCCAGTTTGTGTGTCATCCGAAAGAGGGCTTGACCATTAAGAC 723
DB 601 CTGCTCCCAACCCAGTTTGTGTGTCATCCGAAAGAGGGCTTGACCATTAAGAC 660
QY 724 ATCACTAAGCAGACCCAGTCCCGGATGATATCCATGAAAAAGAACTGAGAGCTGCA 783
DB 661 ATCAACCAACAGACCCAGTCCCGGATGATATCCATGAAAAAGAACTGAGAGCTGCT 720
QY 784 GAGAAAGCTGTGACCATCCATCCAGCCAGAGGGGACTTGTGAAGCATGCCGATGAT 843
DB 721 GAGAAAGCTGTGACCATCCATCCAGCCAGAGGGGACTTGTGAAGCATGCCGATGAT 780
QY 844 CTGGAATCAATGACAGAAAGAGGAGATGAGACCAAACTAGCCGAAAGATTCCTGAAA 903
DB 781 CTGGAATCAATGACAGAAAGAGGAGATGAGACCAAACTAGCCGAAAGATTCCTGAAA 840
QY 904 ATCTTGCAACAATGAGCTGTGTGTAAGACTGATGTAAGAAAGAGGCAAAATTTGAG 963
DB 841 ATTTTACTCATATATACTTTGTTGAGCGTCTTATGTTAAAGAGAAATCTTAAA 900
QY 964 AAAATTGAACATGAAAGAGGAGCAAGATTAACATCTCATCTTGGCAGATTTGACATA 1023
DB 901 AAAATTGAACATGAAAGAGGAGCAAGATTAACATCTCATCTTGGCAGATTTGACATA 960
QY 1024 TACAACCCGAGAAAGAACATCACTGTGAAGGGCAGACAGTGGCCAGTGTGAG 1083
DB 961 TATATTCAGAGACCACTATTTACGTTAAAGGCAATGTTGAGACATGTGCCAAAGCTGAG 1020
QY 1084 ATGAGATTTAAGAAAGCTGTGAGGCTTTGAAATGATGCTGGCTTTACACC 1143
DB 1021 GAGGAGATCATGAAGAAATCAAGGAGCTTTATGAAATGATGCTTTATGATCTT 1080
QY 1144 CACTCCGGAT-----ACTTCTCA 1162
DB 1081 CAGGACATTTAATCTGTGATTAATCTGAACGCTTGGGTCTGTCCACCACTTCA 1140
QY 1163 GCCGTACCCCATCAACGATTTGGCCGTTCCCGCATCACTC----- 1208
DB 1141 GGGATGCTACCTCCCACTCAGGGCCCCCTTACGCACTGACTCTTCCCTTACCCGAGTTT 1200
QY 1209 -TTATCAGAGCAGAGATTTGATCTCTTCATCCCAACCAAGGCTGTGGGCCATC 1266
DB 1201 GAGCAATCAAGAAAGGAGACTGTTCATCTGTTATCCAGCTCATCAGTGGGCTATC 1260
QY 1267 ATGCGAAGAAAGGGGAGCAGATTCATCAACAGCTGGCGGATTTGCCGAGCTCTATCAAG 1326
DB 1261 ATGCGAAGAAAGGGGAGCAGATTCATCAACAGCTGGCGGATTTGCCGAGCTCTATCAAG 1320
QY 1327 ATTGCCCTGCGAAGGCTCAGACGTCAAGCAAAAGATGTCATCAACCGGGCCACCG 1386
DB 1321 ATTGCTCCAGCGAAGACCAAGATGCTAAAGTGAGATGATTAATCACTGGAACACA 1380
QY 1387 GAAAGCCAGTTCAAGGCCAGGAGCGGATCTTTGGAAACTGAAAGAGAAACTTCTTT 1446
DB 1381 GAGGCTCAAGTTCAAGGCTCAGGAAAGATTTATGAAATAATTAAGAGAAACTTTGTT 1440
QY 1447 AACCCCAAGAAAGATGAGAGCTGGAAGCCATATCAGAGTGCCTTCCACAGCTGGC 1506
DB 1441 AGTCTTAAAGAGAGGTAAGCTTGAAGCTCATATCAGAGTGCATCTTGTGCTGGC 1500
QY 1507 CGGCTGATTTGGCAAAAGGTGGCAAGACCGTGAACGAACCTGAGAACTTAACAGTGCAGAA 1566

DB 1501 AGATTTATGAAAGAGAGGAAACGGTAATTAACCTTCAGAAATTTGTCAAGTGCAGAA 1560
QY 1567 GTCACTGTGCTGTGACCAAAACGCAAGTGAATAAGAGATGATGTCAGAAATTAATC 1626
DB 1561 GTTGTGTGCTGTGACCAAAACGCAAGTGAATAAGAGATGATGTCAGAAATTAATC 1620
QY 1627 GGGCATCTTCTTGTAGCCAGACTGCAAGCGCAAGATCAAGGAAATTTGTACAAAGGTG 1686
DB 1621 GGTCACTTCTTGTAGCCAGACTGCAAGCGCAAGATCAAGGAAATTTGTACAAAGGTG 1680
QY 1687 AAGCAGAGAGAGAGAAATACCTCAGGGAGTGGCTCAG 1726
DB 1681 AAGCAGAGAGAGAGAAATACCTCAGGGAGTGGCTCAG 1720

Search completed: July 24, 2004, 22:29:09
Job time : 1427.28 secs

This Page Blank (uspto)

REFERENCE
AUTHORS
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoke, S., Sasaki, N., Carrinci, P., Komoto, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
RIKEN Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
MEDLINE
PUBMED
20530913
11076861

TITLE
4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE
AUTHORS
5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE
6 (bases 1 to 3592)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Komuro, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takeku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toy, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submision
Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES
source
1. 3592
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB: C330012H03"
/db_xref="MGI: 2416441"
/db_xref="taxon: 10090"
/clone="C330012H03"
/cell_type="ES cells"
/clone_1ib="RIKEN full-length enriched mouse cDNA library"
34..1811
/note="HEPATOCYLLULAR CARCINOMA AUTOANTIGEN homolog [Homo sapiens] (SPTRI|O9YCM1, evidence: FASTY, 88.3%id, 100%length, match=1775)
putative"

ORIGIN
Query Match 58.9%; Score 1935.2; DB 11; Length 3592;
Best Local Similarity 79.7%; Pred. No. 0;

Matches 2720; Conservative 0; Mismatches 469; Indels 224; Gaps 28;

QY	38	GGGGGACCCCGGGCTCTCGGGGAAGAGACGATGATGAACAAGCTTTATACATCGGGAAC	97
Db	3	GGGGGACCCCGGGCTCTC-GGGAAAGCGGAGATGAACAACTGTATACATTTGGGAAC	61
QY	98	TGAGCCCCCGGCGTCAACCGGACGACACTCCGGAGCTTTTGGGAAAGAAAGCTGCCCC	157
Db	62	TGAGCCCCCGGCGTCAACCGGACGACACTCCGGAGCTTTTGGGAAAGAAAGCTGCCCC	121
QY	158	TGGCGGGAAGAGTCTCTGTGAAGTCGGCTTACGCGCTTCGTGAGTACATCCCGACAGAACT	217
Db	122	TGGCGGGAAGAGTCTCTGTGAAGTCGGCTTACGCGCTTCGTGAGTACATCCCGACAGAACT	181
QY	218	GGGCAATCCCGGCGCATGAGACCTCTCGGGGTAAAGTGAATTCATGGGAAATTCATGG	277
Db	182	GGGCAATCCCGGCGCATGAGACCTCTCGGGGTAAAGTGAATTCATGGGAAATTCATGG	241
QY	278	AAGTTGATTAATCTCACTCTTAAAAAGCTTAAGAGACGAGAAATTCAGATTGAAACATCC	337
Db	242	AAGTTGATTAATCTCACTCTTAAAAAGCTTAAGAGACGAGAAATTCAGATTGAAACATCC	301
QY	338	CTCTCACTTCACTGAGTGGAGTGTGGATGAGACTTTTGGCTCAATATGGGACAGTGAAGA	397
Db	302	CGCTCACTTCACTGAGTGGAGTGTGGATGAGACTTTTGGCTCAATATGGGACAGTGAAGA	361
QY	398	ATGTGAACAAGTCAACACACAGACAGAAACCGCGTTGTCAAGCTCAATATGCAACAA	457
Db	362	ACGTGAAGCAAGTCAACACAGATACAGAAATTCAGCTGTCAAGCTCAATATGCAACAA	421
QY	458	GAGAAAGAGCAAAATATGACCAATGAGAAAGCTTAAGCGGCAATCAGTTTGAAGACTCTCT	517
Db	422	GAGAAAGAGCAAAAGTATGAGAAAGCTTAAGCGGCAATCAGTTTGAAGACTCTCTCT	481
QY	518	TCAAGATTCTTCAATCCCGATGAAGAGTGAAGTCCCGCTCGCCCTCAGCAGAGCC	577
Db	482	TCAAGATTCTTCAATCCCGATGAAGAGTGAAGTCTTCTTCAACCCCTCATTCG-----	536
QY	578	AGCGTGAGGACCACTCTTCCGGGAGCAAGGCAAGCCCGCTTGGGGGCACTTCTCAGGCCA	637
Db	537	-----TGCCGGGGAACAAGGCAAGGCAAGCCCGGAGCTTCTTCTCAGGCCA	580
QY	638	GACAGATTGATTTCCCGTCCGGATCCCTGCTGCTCCCAACCAAGTTTGTGTGCAATCATG	697
Db	581	GACAGATTGATTTCCCGTCCGGATCCCTGCTGCTCCCAACCAAGTTTGTGTGCAATCATG	640
QY	658	GAAAGAGGCTTGAACCAATTAAGAAATCACTAAGACAGCCAGTCCCGGTAGATATCC	757
Db	641	GAAAGAGGCTTGAACCAATTAAGAAATCACTAAGACAGCCAGTCCCGGTAGATATCC	700
QY	758	ATGAAAGAGAACTCTGAGCTGCAAGAACTGTCAACATTCATGCAACCCCAAGAG	817
Db	701	ACGAAAGAGAACTCTGAGCTGCAAGAACTGTCAACATTCATGCAACCCCAAGAG	760
QY	818	GGACTTTGAGACATGCGCGATGATTTCTTGAATCATGAGAAAGGAGATGAGACCA	877
Db	761	GGACATTTGAGACATGCGCGATGATTTCTTGAATCATGAGAAAGGAGATGAGACCA	820
QY	878	AACCTGCGGAGAGAAATTCCTGAAATCTTGGCAACATGCGCTTGTGGAAGACTGA	937
Db	821	AACCTGCGGAGAGAGAAATTCCTGAAATCTTGGCAACATGCGCTTGTGGAAGACTGA	880
QY	938	TTGGCAAAAGAGCAAAATTTGAAGAAATTTGAACATGAACAGGACCAAGATTAACA	997
Db	881	TTGGCAAAAGAGCAAAATTTGAAGAAATTTGAACATGAACAGGACCAAGATTAACA	940
QY	998	TTCATCTTTGAGAGATTTGAGATATCAACCCGAAAGAAACCATCAGCTGTGAAGGCA	1057
Db	941	TTCATCTTTGAGAGATTTGAGATATTAACCCGAAAGAAACCATCAGCTGTGAAGGCA	1000
QY	1058	CAGTTGAGGCTGTGCAAGTGTGAGATGAGATTAAGAAAGCTGCGAGGCTTTG	1117
Db	1001	CAATTGAAGCTGTGCAATGCTGAGATTAAGATTAAGAAAGCTTCCGAGAGGCTTTG	1060

```
QY 1118 AAAATGATATGTGCTGTTAAC----- 1140
Db 1061 AGAAGCATGTGCTGCCGTTAACAAAGCAATCTGATCCAGGGCTAAACTCACTAGT 1120
QY 1141 ----- 1140
Db 1121 CACTTGGCATCTTTTGAAGTGTGTGTGCTTCTTCAAGCAGAGGGCCCCGTGAGT 1180
QY 1141 ----- ACCCACTCCGGAGATCTTCTCAAGCTCTGA 1169
Db 1181 TCCCCCGAGTCTCCCTATCAACCCCTTTGCTTACCCACTCCGGATATCTTCCAGCTCTGA 1240
QY 1170 CCCCCATCACCAGTTTGGCCCGCTTCCGSCATCATCACTCTTATCCAGAGCAGAGATTGT 1229
Db 1241 CCGCATACCATTTCCGCCCATTTCCCATCATCATCTCTTACCCAGAGCAGAGAGATGT 1300
QY 1230 GAATCTCTTCAATCCCAACCCAGGCTGTGGGGCCATCATCCGGAGAGAGGGGGCACAAT 1289
Db 1301 AAGTCTCTTCAATCCCAACCCAGGCTGTGGGTGTATCATCGGAGAGAGGGGGCACAAT 1360
QY 1290 CAAACAGCTGGCGAGATTCGCGGAGGCTCTATCAAGATTGCCCCGCGAGAGCCCA 1349
Db 1361 CAAACAGCTGGCTGATTTTGTGTGTCTTCATCAAGATTGCTCCAGAGAGAGGTCCAGA 1420
QY 1350 CGTCAGCGAAAGATGTGTATCATCACCGGGCCACCGGAAGCCAGTTCAAGGCCCAAGG 1409
Db 1421 TGTCAGTGAAGAGATGTGTATCATCATGTGTCTCTGGAAGCCCAAGTTTAAGGCTCAAGG 1480
QY 1410 ACCGATCTTTGGGAACTGAAAGAGAAAACTTCTTTAACCCCAAGAGAGATGAACT 1469
Db 1481 ACCGATCTTTGGGAACTGAAAGAGAAAACTTCTTTAATCCCAAGAGAGATGAACT 1540
QY 1470 GGAAGGCGCATATGAGATGCCCTCTTCCACAGCTGGCCGGGTGATTGGCAAGGTGCCAA 1529
Db 1541 GGAAGGCGCATATCGAGTCCCTCTGTGACCGCTGGCGGTGATTGGCAAGGCGGAAA 1600
QY 1530 GACCGTGAAGGAACTGAGAACTTAAACAGTGCAGAGTATGCTGTGTGACCAAC 1589
Db 1601 AACCGTGAAGAGGTGAGAACTTGAAGAGTGCAGAACTTATGCTGTGTGACCAAC 1660
QY 1590 GCCAGATGAAGATGAGAACTGATGTCAGAAATTATCGGCACTTTTGTGTAAGCCAGAC 1649
Db 1661 GCCAGATGAAGATGAGAACTGATGTCAGAAATTATCGGCACTTTTGTGTAAGCCAGAC 1720
QY 1650 TGCAAGCGCAAGATGAGGAAATTGTAAACAAGGTGAAGACAGCAGAGCAAGAAATGCC 1709
Db 1721 TGCAAGCGCAAGATGAGGAAATTGTAAACAAGGTGAAGACAGCAGAGCAAGATGCC 1780
QY 1710 TCAGGAGTCTGCTCAAGCGCAGCAAGTAGGGCTCCACAGGCAACGCAAAACAAG 1769
Db 1781 TCAGGAGTCTGCTCCACAGCGCAGCAAGTAGGGCTCCACAG--CACACAGCAAGCAACG 1838
QY 1770 ATGAATGTAGCCCTTCCAAACCTGACAGATGAGCAAAACGACAGCAGCAGATCGG 1829
Db 1839 ATGAATGTAGCCCTTCCAAACCTGACAG--ATGAAGCAAAAC--AGCAGCAGATCGG 1894
QY 1830 AGCAAAACCAAGACAT-CTGAGGAATGAGAGTCTGCGGAGCGCGCAGGGACTGTGCC 1888
Db 1895 AGCAAAACCAAGAGCATCCGAGAGAGTGGCAGATCTGCAAGCAGCGAGGCGCTGCG-- 1950
QY 1889 GAGGCCCTGAGAAACCCAGGGGCGAGAGAGGGGCGGAGAAAGTCAGCGAGTTTGCAGAA 1948
Db 1951 -AGACCTTACACATCTTGGGATCCAGAGGGGCAAGGGAG--GCCAGTTTGTCCAG 2006
QY 1949 ACCACCG--AGCCCGCCTTCCGCGCCCAAGGGCTTCTGCAAGCTT--CAGCCATCCACT 2004
Db 2007 AACACCGTGTGGCTGCGCCCGCAGCTTCCCTGTGCTTCTGCAAGCATTAACAGCATCACT 2066
QY 2005 TCACATCAGCTGAGATCTCTCTGAACCTCCACAGAGGCTATGCCCTTTTAACTGAATGA 2064
Db 2067 GCCCATCAACTAGAT-TCTCTCAAGTTCCAGAGCGCTATCCCTTGGTTGAATCA 2125

QY 2065 CATAGTGAAGCTGTTCAAAGCCAAAGCAAAATGCAACACC---TTTTCTGTGCAAAAT 2120
Db 2126 CATAGTGAAGCTGTTCAAAGCCAAAGCAAAATTCCTACCCCTTTCTTGTGTGGAAGAT 2185
QY 2121 CGTCTGTGATCATGTGTGTATCATATTTAGAAAGGGA--GATGTTAAGATATGTGGCTGTG 2179
Db 2186 TGTCTGTGATCATGTGTGTATCATATTCAGAGGGGAAGGATTTTAAAGATGTGGCTGTG 2245
QY 2180 GGTTCACAGGGTCTGTGAGCGGTATATATTTTAAAGATATATATCAATTAATCA 2239
Db 2246 GGTTCACAGGGTCTGTGAGCGGTATATATTTTAAAGATATATATCAATTAATCA 2305
QY 2240 CTAATCCAAATTTTAAATCAATTAATTTTCTTTTAAAGAAAGCAGCTT 2299
Db 2306 CTAACCAATTTTAAATCAATTAATTTTCTTTTAAAGAAAGCAGCTT 2365
QY 2300 TTCTAGACTTTAAAGATTAAGTCTTGG--GAGGTCTAGGGTGAAGAGAGCTTGA 2358
Db 2366 TTCTAGACTTTCAAGAAATTAAGTATTGGGTAGGTCCTGTATTAAGAAAGAACTTGA 2425
QY 2359 GGC-CACCCGCAAAAATTCAACCGAGGGAATCTGTGGAAGGACACTCAGGGCAGT 2417
Db 2426 GCCACCCACAAAAAATTCACTTGAAGGGAATCTGTG-C-AAGGACCTTAATGGCAGT 2484
QY 2418 TCTGGATCACCTGTGTATGTCAACAGAGGATACCGTCTCTTGAAGAGGAACTCTGT 2477
Db 2485 TGTGGGTACCTGTGTGTGTCAACAGAGGATACCGTCTCTGAGAGGGCCCTAGGCTC 2544
QY 2478 CACTCTCATGSCCTGTACTACTCATACCACTTCTCT-----T 2517
Db 2545 CTGGTCACTGTCTTCTCCACACACCACTGTCTCATGTTTCAAGGTTTAACTGA 2604
QY 2518 TGCTTCAAGGTTTAAACTGTTTTTGTGATCTGCTATATATATCTGTCTCTCT 2577
Db 2605 TTTTGGTGTGTTTGTGTTTGTGTTTGTGATCTGTATATAGTCTGTCTCTC-- 2662
QY 2578 GTTATGTCTCCCTCCCTCCCTCCCTCCCTCTCTCTCATCTCAATCTTTTGAATTTCT 2637
Db 2663 ----TATCTCCCTCTCAACCTTCTCTGTCTTCTTACCAAAATCTTTTGAAGTCCCT 2718
QY 2638 CATCCCTCATCTCAATCCCGTATCTAGCAGCC---CGCCCGCCCGCAAGCAAGCTG 2694
Db 2719 TCGTCTTTTCTGTCTCCCTGTATCTATGTGACCCCTCCCTCCCGCAAGCAAGCTG 2778
QY 2695 CTCTGATATCATATCAACAAAGAAAGAAAGCGAAACACACAAACGACCTCACTT 2754
Db 2779 CTCTGATATCATATCAACAAAGAAAGAAAGCGAAACACACAGCAGCTCAGCTT 2838
QY 2755 ACACTTGTATCTCAAAAGAAAGAGTCAATGTACTTGTCTTACGCTTTTGAAGAG 2814
Db 2839 ACACTGTTTACTCAAGCGAAAGAGTGGCGGTACTTGTCTTACG-ATTGGAAGAG 2897
QY 2815 AAAACAGAAACCCACCAACCAACCAATCAACCAACCAAGAAATTCACATTA 2874
Db 2898 AAAAGCAGATCCAGCAAAATTAACAGCCAGCCAAACCAAG--AAAATTCACAAAGAA 2956
QY 2875 GAAATGATTTGTCTTTTGCATTTTGTGTATTAAGCAATATTCAGCAAAATGAT 2934
Db 2957 GAATGGA---CTCTCTTCATCTTGTATTAACCAATCAATTTTGTGAATTTCTT 3012
QY 2935 CTTTTCTTAAAAAAAATGTGAGAGAAAGTAAATTTTCAAGGTTTGGCCAG 2994
Db 3013 TCTTTCTTAAAAATTTAAAGTGAAGAAATGCAAAATTTATACAGGTTTGGCCAG 3072
QY 2995 GGGTTAAATTCACAGATTTTAAAGGAAACACACAG--AGAAGCTACCTCAG 3051
Db 3073 GGGTTAAATTTACAGATTTTAAAGGAAACACACAAATTAAGCTACCTCAG 3132
QY 3052 GTG-TTTTAACTCAGACCTTGTCTGTGTGTTTCCCTTAAGATTTGTTAAAGCTGATA 3110
Db 3133 GTGTTTAACTCAGACCTTGTCTGTGTGTTTCCCTTAAGATTTGTTAAAGCTGATA 3192
QY 3111 GTTGAAGCATTTTATTTTAAATTAATTAATGAGTGGAAAAAATTAAGATATCAAC 3170
```

Db 3193 GTTGAGACA-TTTTATTTTAAATAATAGTT-GAAAAAATAATGATGAC 3250
QY 3171 TGGCAGCTGGAGAGGTGACAGTCCAAAGTGTGCAAGCTGTTCTGAATGCTTCGCG 3230
Db 3251 TGGCAGCTGGAGAGGTGACAGTCCAAAGTGTGCAAGCTGTTCTGAATGCTTCGCG 3305
QY 3231 TAGCCAAAGACCAATATGAGCTTCTTTTGGACAAACCTTGAAATGTTTATTT 3283
Db 3306 TAGCCAAAGACCTGAGTGGCTTCTTTGGACAAACCTTGAAATGTTTATTT 3358
RESULT 2
AK077404 3262 bp mRNA linear HTC 18-SEP-2003
LOCUS Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched
DEFINITION library, clone:5430439D08 product:HEPATOCELLULAR CARCINOMA
AK077404.1 GI:26097330
ACCESSION AK077404.1 GI:26097330
VERSION HTG; CAP trapper.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE Itch, M., Komo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS Shibata, K., Itch, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
TITLE Komo, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itch, M.,
Sunt, N., Ishi, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujimake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
TITLE FANTOM Consortium.
JOURNAL Functional annotation of a full-length mouse cDNA collection
AUTHORS Nature 409, 685-690 (2001)
TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research
JOURNAL Group Phase I & II Team.
REFERENCE Analysis of the mouse transcriptome based on functional annotation
AUTHORS of 60,770 full-length cDNAs
TITLE Nature 420, 563-573 (2002)
JOURNAL Nature 420, 563-573 (2002)
REFERENCE Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
TITLE Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M.,
AUTHORS Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itch, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, U., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sozabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki. The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
FEATURES
source
1. .3262
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:5430439D08"
/db_xref="MG1:2393565"
/db_xref="Laxson:10090"
/clone="5430439D08"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="6 days neonate"
misc_feature
16. .1796
/note="HEPATOCELLULAR CARCINOMA AUTOANTIGEN homolog [Homo
sapiens] (SPTR1096M1, evidence: FASST, 88.3%id,
100%length, match=1775)
putative. 3249. .3254
/note="putative"
3262
polyA_site
/note="putative"
ORIGIN
Query Match 53.8%; Score 1765.4; DB 11; Length 3262;
Best Local Similarity 78.1%; Pred. No. 0;
Matches 2581; Conservative 0; Mismatches 466; Indels 256; Gaps 28;
QY 58 GGAAGAGACGAGATGATGAAAGCTTTTACATCGGGAACCTGAGCCCGCTCACCGCC 117
Db 4 GGAAGAGACGAGATGATGAAAGCTTTTACATCGGGAACCTGAGCCCGCTCACCGCC 63
QY 118 GACGACCTCCGCGAGCTCTTTGGGGACAGAAAGCTGCCCTGCGGACAGAGTCTGCTG 177
Db 64 GACGACCTCCGCGAGCTCTTTGGGGACAGAAAGCTGCCCTGCGGACAGAGTCTGCTG 123
QY 178 AAGTCGGCTACGCTTCTGTCGATCCCGACCGAACTGGGGCATTCGGCGCATCGAG 237
Db 124 AAGTCGGCTACGCTTCTGTCGATCCCGACCGAACTGGGGCATTCGGCGCATCGAG 183
QY 238 ACCCTCGGGGTAAAGTGAATTCGATGGGAAATCATGGAAGTTGATTACTCACTCT 297
Db 184 ACCCTCGGGGTAAAGTGAATTCGATGGGAAATCATGGAAGTTGATTACTCACTCT 243
QY 298 AAAAAGCTAAGAGCAGAAATTCGATTCGAAATCCCTCTCACTGCGAGTGGAG 357
Db 244 AAAAAGCTAAGAGCAGAAATTCGATTCGAAATCCCTCTCACTGCGAGTGGAG 303
QY 358 GTTGTGATGAGCTT-TTGGCTCAATATGAGACAGTGGGAATGTGGAACAATCAAC 416
Db 304 GTTGTGATGAGCTT-TTGGCTCAATATGAGACAGTGGGAATGTGGAACAATCAAC 363
QY 417 AGACACAGAAACGCGCTTCAACGTCACATATGCAACAAGAGAAAGCAAAATAGC 476
Db 364 AGATACAGAAACGCGCTTCAACGTCACATATGCAACAAGAGAAAGCAAAATAGC 423

QY	477	CATGGAAGACTAAGGGGCACTAGTTTGAACTACTCTTAAATTTCTTACTACCTCC	536
Db	424	TATTGGAAGCTAAGGGGCAAGTTTGAGGCTACTCTTCAAAATTTCTTACTACCTCC	483
QY	537	GGATGGAAGAGTGAAGCTCCCTTCGGCCCCGAGGAGGCCAGCTGGGGACCACTCTTC	596
Db	484	CGATGGAAGAGTGAAGCTCTCTTCCACCCCTCAATCG-----TGC	522
QY	597	CCGGAGCAAGGCGCAAGCCCTGGGGGCACTTCTCAGGCGACAGATGATTTCCGCT	656
Db	523	CCGGAAACAAGGCGAAGGCCCGGAGCTCTTCTCAGGCGACAGATGATTTCCGCT	582
QY	657	GGGATCTCTGTCTCCCAACCAAGTTTGTGTGCATCATCGAAAGAGGCTTGACAT	716
Db	583	GGGATCTCTGTCTCCCAACCAAGTTTGTGTGCATCATCGAAAGAGGCTTGACAT	642
QY	717	AAAGAACATCACTAAGCAACCAAGTCCCGGGTAAGTATCCATAGAAAAGAACTCTGG	776
Db	643	AAAGAACATCACTAAGCAACCAAGTCCCGGGTAAGTATCCATAGAAAAGAACTCTGG	702
QY	777	AGCTCAGAGAAGCTGTCAACATCCATCCACCCAGAGGGGACCTTCTGAAGCAATGCC	836
Db	703	GGCTCAGAGAAGCTGTCAACATCCATCCACCCAGAGGGACATCTGAAGCAATGCC	762
QY	837	CATGATTTCTGAATCATGCAAGAAAGGCAAGATAGACAAACTAGCCGAGAATTTCC	896
Db	763	CATGATTTCTGAATCATGCAAGAAAGGCAAGATAGACAAACTAGCTGAAGAGTTCC	822
QY	897	TCTGAAATCTTGGGACCAATGGCTTGGTTGGAAAGCTGATTTGAAAAGAAAGGCAAAA	956
Db	823	TCTGAAATCTTGGGACCAATGGCTTGGTTGGAAAGCTGATTTGCAAAAGGCAAAA	882
QY	957	TTTGAAGAAAATTGAACATGAAACAGGGACCAAGATTAACATCTCATCTTGGACAGATT	1016
Db	883	CCTGAAGAAAATTGAACATGAGACAGGGACCAAGATTAACATCTCATCTTGCAGAGATT	942
QY	1017	GAGCATATPACACCCGGAAGAACCATCACTGTGAAGGCACAATTGAGGCTGTG-CCA	1075
Db	943	GAGCATTTATPAAACCCGAGAAACCATCACTGTGAAGGCAACATTGAAGCTGTGCCCA	1002
QY	1076	GTGCTGAATTAAGATTATGAAGAGCTGCTGAGGCTTTGAAATATATGTGAGCTG	1135
Db	1003	ATGCTGAATTAAGATTATGAAGAGCTGCCAGAGGCTTTGAGAAACACATGTGCGCG	1062
QY	1136	TTAAC-----	1140
Db	1063	TTAACCAACAGCCATCTGATCCAGGGCTPAAACCTCAGTGCACCTTGTTCGA	1122
QY	1141	-----	1140
Db	1123	CTGACCTGTGTGCTCTCTCCACAGAGAGGCCCGCTGGAGTTCCCCCAAGTCTCTCT	1182
QY	1141	-----ACCACTTCGGATATCTTCTCCAGCTGTATCCCTCCATCACAATTGG	1186
Db	1183	ATCAACCCCTTGTACCAACCACTCCGATCTCTCCAGCTGTATCCCTCATCACCAATTGG	1242
QY	1187	GGCGTTCCCGGATCATCACTCTTAACAGAGGCAAGATTTGATTTCTTCACTCCCA	1246
Db	1243	GGCAATTCACATCATATCACTCTTACCCAGAGGCAACATGTAAAGTCTCTTCACTCCCA	1302
QY	1247	CCAGAGCTGTGGGCGCATCATCGGGAAGAGGGGSCACATCAAAACAGCTGCGAGAT	1306
Db	1303	CCAGAGGCTGTGGGCTATCATTCGGGAAGAGGGGSCACATCAAAACAGCTGCTGAT	1362
QY	1307	TGCGCGAGCTGTATCAAGATTGCCCTGTGGAAAGCCAGACGTCAAGCAAAAGATGG	1366
Db	1363	TTGCTGTGCTCCATCAAAAGATTGCTCAGAGAGGTCAGATGTCAGTGAAGGATGG	1422
QY	1367	TCATCATACCGGGGCAACCGGAACCCAGTTCACAGGCCAGAGACGATCTTTGGAAAC	1426
Db	1423	TCATCATACTGTGCTCTCTGAAGCCAGATTAAAGCTCAAGAGCGATCTTTGGAAAC	1482

QY	1427	TGAAAGAGGAAA	CTCTTTTAA	CCCCAAGAAGATGA	ACTGGAGCCATAT	CAGAG	1468		
Db	1483	TTAAAGAAAGAAA	CTTTTAT	TCCAAAAGAAAGTGA	MACTGGAGCCATAT	CAGAG	1542		
QY	1487	TGCCCTCTCC	CACAGCTG	GCAGGAGT	TGGCAAGGTGGGCAAGCCGTGAAC	CACTGC	1546		
Db	1543	TCCCTCTGTC	CACCGCTGG	CCGGTGA	TGGCAAGGTGGGCAAGCCGTGAAC	CACTGC	1602		
QY	1547	AGAACTTAA	CCAGTCCAGAA	GTCACTGTGCTG	CAACACGCACATGA	AAATGAG	1606		
Db	1603	AGAACTTAA	CCAGTCCAGAA	GTCACTGTGCTG	CAACACGCACATGA	AAATGAG	1662		
QY	1607	AGTATGCTG	CAGAA	TTATCGG	GCACCTTTGCTAG	CCAGCTGCACAGCGCAAG	ATCA	1666	
Db	1663	AGTATGCTG	CAGAA	TTATCGG	GCACCTTTGCTAG	CCAGCTGCACAGCGCAAG	ATCA	1722	
QY	1667	GGGAATATG	TACAA	CAGGTGAA	AGCAGCAGAGCA	AAATACCTCAGG	GAATCCGCTTCAC	1726	
Db	1723	GGGAATATG	TACAA	CAGGTGAA	AGCAGCAGAGCA	AAATACCTCAGG	GAATCCGCTTCAC	1782	
QY	1727	AGCGCAGCA	GTGAGGCTCC	CAACAGGCA	CCAGCAAA	CAACGATGA	TGAGCCCTTCC	1786	
Db	1783	AGCGCAGCA	GTGAGGCTCC	CAACAGGCA	CCAGCAAA	CAACGATGA	TGAGCCCTTCC	1840	
QY	1787	AACACTTAC	AGATGAG	ATCCAA	CGCAGCCAG	CCAGTGGG	AGCAACCAAGACAT	1846	
Db	1841	AACACTTAC	AGATGAG	ATCCAA	CGCAGCCAG	CCAGTGGG	AGCAACCAAGACAT	1896	
QY	1847	-CTGAGAA	TAGAA	GTCTGCG	AGGCGCCAGG	GAATCTG	CCGAGGCCCTG	AGAACCC	1905
Db	1897	-CTGAGAA	TAGAA	GTCTGCG	AGGCGCCAGG	GAATCTG	CCGAGGCCCTG	AGAACCC	1952
QY	1906	AGGCGCCAGG	AGGCGCGG	GAAGATGA	GCAGGCTTTG	CCAGAAC	CAACCG--AGCCCGC	1963	
Db	1953	TGGGATTA	GAGAGGG	CACAGGGA	G---GCCAG	TTTCCAGAA	CAACCGCTTGCCTGC	2009	
QY	1964	CTCCGCC	CCCCCAGGG	CTTCTG	CAGGCTT--CAGCC	ATCACTTCA	CCATCCATCGGAT	2021	
Db	2010	CTCCGCC	CCCCCAGGG	CTTCTG	CAGGCTT--CAGCC	ATCACTTCA	CCATCCATCGGAT	2069	
QY	2022	CTCTCCGAA	TCCCA	CGAGCTAT	CCCTTTAG	TGAATTA	CTATAGGTGA--CGTGT	2079	
Db	2070	-TCTCTCAG	TTCCCA	CGAGCTAT	CCCTTTAG	TGAATTA	CTATAGGTGA--CGTGT	2128	
QY	2080	TCAAAGCC	CAAGCA	AAATGCA	CACTTTTTC-----TGTGG	CAAAATGCTCTG	TGATCATG	2134	
Db	2129	CAAAAGCC	CAAGCA	AAATGCTT	CTAGCCCTTTCTTTG	TGTGA	AAAGTTGTTCTCTG	2188	
QY	2135	TGTGTCA	TATTTAGAA	AGGAA--GATG	TTAAGATATG	GGCCCTGTGG	TATTA	2193	
Db	2189	TATGTCA	TATTTAGAA	AGGAA--GATG	TTAAGATATG	GGCCCTGTGG	TATTA	2248	
QY	2194	CTGCGAC	CGGTAA	TATTTTAA	GAATATTA	CTCAATTA	CTCACTCAATTTT	2253	
Db	2249	CTGCGAC	CGGTAA	TATTTTAA	GAATATTA	CTCAATTA	CTCACTCAATTTT	2308	
QY	2254	TATCAAT	TTAATTTT	TTTTCTTTTAA	AGAAAGAGG	CTTTTCA	GCTTAA	2313	
Db	2309	TATCAAT	TTAATTTT	TTTTCTTTTAA	AGAAAGAGG	CTTTTCA	GCTTAA	2368	
QY	2314	GAATAA	GTCTTTGG	AGGTCTC	ACGT-----GTAGA	AGAGCTTTG	AGGCCACC	2365	
Db	2369	GAATAA	GTCTTTGG	AGGTCTC	ACGT-----GTAGA	AGAGCTTTG	AGGCCACC	2428	
QY	2366	CGCA	CAAA-----TTCA	CCCAAGG	CAAAATTTTTC	CAAGGCA	CACTTAA	2401	
Db	2429	CGCA	CAAA-----TTCA	CCCAAGG	CAAAATTTTTC	CAAGGCA	CACTTAA	2488	
QY	2402	GGAAC	ACTCAGG	CACTGTG	ATCACTGTG	TATCA	CAAGAGGATAT	2461	
Db	2489	AGTTTGT	GGGTTCC	ACCTGTTTCT	TTTTCAG	AGCAAGCA	AAAGGAAACCTTCCCT	2548	
QY	2462	GAAAG	AGAA	CTCTGTCA	CTCTCATG	CTGTGTAG	CTATACACCATTTCT	2521	

misc_feature
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="9.5 days embryo"
1..2964
/note="HEPATOCELLULAR CARCINOMA AUTOANTIGEN homolog [Homo sapiens] (SPTR|Q9Y6M1, evidence: PASTY, 88.3%id, 100%length, match=1775)"

Query Match 49.4%; Score 1622; DB 11; Length 2964;
Best Local Similarity 78.2%; Pred. No. 4.2e-307;
Matches 2329; Conservative 0; Mismatches 435; Indels 216; Gaps 23;

```
QY 247 GGTAAAGTGAATTCATGCGAATAATCATGAAAGTTGATTAAGTCTCTTAAAGCTA 306
Db 30 GGTAAAGTGAATTCATGCGAATAATCATGAAAGTTGATTAAGTCTCTTAAAGCTA 89
QY 307 AGGAGCGAAGAAATTCATGCTGAAACATCTCTCTCTCACTGCAAGTGGAGGTGTGAT 366
Db 90 AGGAGCGAAGAAATTCATGCTGAAACATCTCTCTCTCACTGCAAGTGGAGGTGTGAT 149
QY 367 GSACTTTTGGCTCAATATGAGGAGTGGAGAAATGAGCAAGTCAACAGACAGACAA 426
Db 150 GGGCTGTGTGGTGAATATGGACAGTGGAGAACTGGAGCAAGTCAACAGATACAGAA 209
QY 427 ACCGCGCTTGCAACGTCAATATGCAACAAGAGAGCAAAATATAGCCATGAGAG 486
Db 210 ACTGCCCTGTGCAACGTCACTATATGACAAAGAGAGGCAAGCTATAGATGAGAG 269
QY 487 CTAAAGCGGAGATGCTTTGAGAACTACTCTTCAAGATTCTTCAATCCCGAGTAAAG 546
Db 270 CTAAAGCGGAGATGCTTTGAGAACTACTCTTCAAGATTCTTCAATCCCGAGTAAAG 329
QY 547 GTGAGCTCCCTTGGCCCTCAGCGAGCCAGGCTGGGAGCACTCTTCCCGGAGCA 606
Db 330 GTGAGCTCCCTTGGCCCTCAGCGAGCCAGGCTGGGAGCACTCTTCCCGGAGCA 368
QY 607 GGCACAGCCCTGCGGGGCACTTCTCAGCGCAGACAGATTATTTCCGCTCGGATCTG 666
Db 369 GGCACAGCGCCCGGAGACTCTTCTCAGCGCAGACAGATTATTTCCGCTCGGATCTG 428
QY 667 GTCCCAACCAATTTGTTGGTGCATATGAGAAAGAGGAGGCTTACCATTAAGAACATC 726
Db 429 GTCCCAACCAATTTGTTGGTGCATATGAGAAAGAGGAGGCTTACCATTAAGAACATC 488
QY 727 ACTAAGAGACCCAGTCCGAGGTAGATATCATTAAGAAAGAACTCTGAGCTGACAG 786
Db 489 ACTAAGAGACCCAGTCCGAGGTAGATATCATTAAGAAAGAACTCTGAGCTGACAG 548
QY 787 AAGCTGTCAACATCCATGCAACCCAGAGGAGCTTCTGAAACATGCGCATGATCTT 846
Db 549 AAGCTGTCAACATCCATGCAACCCAGAGGAGCATCTGAAACATGCGCATGATCTT 608
QY 847 GAAATCATGCAAGAAAGGCGAGATGAGACCAACTAGACCGAAGATCTCTGAAATC 906
Db 609 GAAATCATGCAAGAAAGGCGAGATGAGACCAACTAGACCGAAGATCTCTGAAATC 668
QY 907 TTGGCACAACATGCTTGGTTGGAAGCTGATTGAAAGAGAGGCAAAATTGGAAGAA 966
Db 669 CTGGCCCAACATGCTTGGTTGGAAGCTGATTGGAAGAGGCAAACTGGAAGAA 728
QY 967 ATTGAACATGAAACAGGAGCAAGATTAACATCTCATCTTTGAGAGATTGACATATAC 1026
Db 729 ATTGAACATGAAACAGGAGCAAGATTAACATCTCATCTTTGAGAGATTGACATATAC 788
QY 1027 AACCCGGAAGAAACATCTGCTGAAAGGCAAGTTGAGGCTGTGCCAGTGTGAGATA 1086
Db 789 AACCCGGAAGAAACATCTGCTGAAAGGCAAGTTGAGGCTGTGCCAGTGTGAGATA 848
QY 1087 GAGATTATGAAGAGCTCGAGAGGCTTTGAAATGATATGCTGCTGTAAAC----- 1140
Db 849 GAGATTATGAAGAGCTCGAGAGGCTTTGAAATGATATGCTGCTGTAAACAA 908
QY 1141 ----- 1140
```

```
Db 909 GCCAATCTGATCCAGGCTTAACCTCAGTGCATCTTGGCATCTTTGACTGACTGCT 968
QY 1141 ----- 1140
Db 969 GTGCTTCTCCACAGCAGGAGCCCGTGGAGTTCCTCCCTCATCACCCCTT 1028
QY 1141 ----ACCCATCTCCGATCTTCTCCAGCTGTACCCCATCAACGTTTGGCCGCTTCCG 1197
Db 1029 GCTAACCACTCCGATCTTCTCCAGCTGTACCCCATCAACGTTTGGCCCATTC 1088
QY 1198 CATCATCACTCTTCCAGAGCAGAGATGTAATCTCTTCAATCCCAACCAAGCTGTG 1257
Db 1089 CATCATCACTCTTCCAGAGCAGAGATGTAATCTCTTCAATCCCAACCAAGCTGTG 1148
QY 1258 GCGCCATCATCTGGGAAGAAAGGAGGCAACATCAACAGCTGGGAGATTTGGCCGAGCC 1317
Db 1149 GGTGCTATCATCTGGGAAGAAAGGAGGCAACATCAACAGCTGGGATTTGGCTGTGCC 1208
QY 1318 TCTATCAAGATGCTCCCTGCGGAAGGCCCAAGCTGCGGAAGAGATGTCATATCAC 1377
Db 1209 TCTATCAAGATGCTCCCTGCGGAAGGCCCAAGCTGCGGAAGATGTCATATCAC 1268
QY 1378 GGGCCACCGGAAAGCCAGTTCAAGGCCCAGGAGCCGATCTTTGGGAAATGAAAGAGAA 1437
Db 1269 GGTCTCTCTGAAGCCAGTTTAAGCTCAAGAGCCGATCTTTGGGAAATGAAAGAGAA 1328
QY 1438 AACTCTTTTAAACCCCAAGAAAGATGAAGCTGGAACGCAATACAGATGCCCTTTC 1497
Db 1329 AACTCTTTTATCCCAAGAAAGATGAAGCTGGAACGCAATACAGATGCCCTTTC 1388
QY 1498 ACGCTGGCCGGGTGATTGGCAAAAGTGGCAAGACGTGAACGAATCTGCAAACTTAAC 1557
Db 1389 ACGCTGGCCGGGTGATTGGCAAAAGTGGCAAGACGTGAACGAATCTGCAAACTTAAC 1448
QY 1558 AGTGAAGAAATCATCTGCTGCTGAGACCAACCCGATGAATGAAGAGATGCTC 1617
Db 1449 AGTGAAGAAATCATCTGCTGCTGAGACCAACCCGATGAATGAAGAGATGCTC 1508
QY 1618 AGAATTATCGGGCACTTCTTTGCTAAGCAGATGCAACAGGCAAGATCAGGAAATTGTA 1677
Db 1509 AGAATTATCGGGCACTTCTTTGCTAAGCAGATGCAACAGGCAAGATCAGGAAATTGTA 1568
QY 1678 CAACAGGTGAAGCAGCAGAGCAAAATCCTTAGGAGTCCCTCAACAGCCGAGAG 1737
Db 1569 CAACAGGTGAAGCAGCAGAGCAAAATCCTTAGGAGTCCCTCAACAGCCGAGAG 1628
QY 1738 TGAGGCTCCCAAGGACCAAGCAAAACGATGAATGAATGAGCCCTTCCAAACCTGACA 1797
Db 1629 TGAGGCTCCCAAGG--CACTAGCAAGCAACGATGAATGAGCCCTTCCAAACCTGACA 1686
QY 1798 GAATGAGCCAAACGAGCCAGCCAGATCGGAGCAAAACCAAGACCAT--CTGAGAGATG 1856
Db 1687 G-ATGAGACCAAC--AGCAGCAATATGGAGCAAAACCAAGAGATCCGAGAGATG 1742
QY 1857 AGAAGTCTGCGAGCGCGCGAGGAGACTCTGCCAGGCTCTGAGAACCCGAGGCGAGG 1916
Db 1743 CGCAGTCTGCAAGCAGCAGGAGGCTGCG--AGACCTCTCAACATCTGGGATCCAGG 1797
QY 1917 AGGAGCGGAGAAAGGTGAGGCAAGTTGCGCAACCAACCG--AGCCCGGCTCCGCCCCC 1974
Db 1798 AGGAGCAAGGAG--GCCAGGTTTCCGAAACCCGCTTGGCTGCCCCCAGCTTCC 1854
QY 1975 CAGGGCTTTCAGAGCTT--CAGCCATCACTTCAACATCACTCGAGATCTCTGTAAC 2032
Db 1855 CCGGCTTTCAGAGGATTAACAGCCATCACTAGCCCATCACTCATGAT--TCTCTCAGT 1913
QY 2033 TCCCAAGAGCTATCTCTTTTGTGTAATTAACATATGTAACGTGTTCAAGCAAGCA 2092
Db 1914 TCCCAAGAGCTATCTCTTTGTTGTAATTAACATATGTAACGTGTTCAAGCAAGCA 1973
QY 2093 AATGACACACC-----TTTTGTGGCAAAATCGTCTGTATCATGTGTGATTTAG 2148
```


Db 1974 AATTCCTGACCGCTTCTTGTGTGGAAGATGTCCTGATCATGATGATCATATCAG 2033
 QY 2149 AAGGAGAA-GATGTTAAGATATGTGGCTGTGGGTTACACAGGGTGCCTGAGGGGAT 2207
 Db 2034 AAGGAGAAAGATGTTAAGAAATGTGGCTGTGGGTTACACAGGGTGCCTGAGGGGAT 2093
 QY 2208 ATATTTTGAAGAAAT 2267
 Db 2094 ATATTTTGAAGAAAT 2153
 QY 2268 TTTTCTTTCTTTCTTTTAAAGAAAGACAGGCTTTTCTAGACTTTTAAAGAAATGCTT 2327
 Db 2154 TTTTCTTTCTTTCTTTTGAAGAAAGACAGGCTTTTCTAGACTTTTAAAGAAATGCTT 2213
 QY 2328 G-GAGGTCACGGGTGTGAGAGAGAGCTTTGAGGC-CACCGGACAAATTTACCCAGAG 2385
 Db 2214 GTGAGTCACCTGGATATGAAAGAACTTTGAGCCACCCACACAAAATTCACCTGAG 2273
 QY 2386 GGAATATCGTGGGAAAGACACCTCAGGAGTGTGATGATCAGCTGTATGTCACAGAA 2445
 Db 2274 GGAATATCTTGT-C-AGGACACTTAATGCACTTGTGGGTCACCTGTCTGTCTGACAGAA 2332
 QY 2446 GGGATACCGTCTCTTGAAGAGAAATCTGTCACTCTCATGCTGTCTAGCTCATACA 2505
 Db 2333 GGGATACCGTCTCTGAGAGGGGCGCTGAGCCCTGTCACCTGTTCTCCACACACCCA 2392
 QY 2506 CCCATTTCTCTT-----TGCTTACAGGTTTAACTGCTTAACTGCTTAACTGCTT 2545
 Db 2393 TTGCTCAGCTTCAACAGGTTTAACTGATGATGATGATGATGATGATGATGATGAT 2452
 QY 2546 GCATCTGCTATATATATATCTGTCTCTCTGTTATCTCTCCCTCCCTCCCTCCCTCC 2605
 Db 2453 GCATCTGCTATATATATGTTCTGTCTCTC-----CTATCTCCCTCCCTAACTCTCT 2506
 QY 2606 TTCTTCTCATCTCATCTTTTGAATTTCTCATCTCTCATCTCATCTCATCTCATCT 2665
 Db 2507 GCTTCTTTTACCAAAATCTTTTGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2566
 QY 2666 GAGC-----CCCCCCCCCAGGCAAGAGAGTCTGATCATCATCATCATCATCATCAT 2722
 Db 2567 GAGCCT 2626
 QY 2723 CAAAAGCAGAAACACACAAACAGCCTCACTTACCTGTTACTCAAAAGAGAGAGT 2782
 Db 2627 CAAATGCGAAGACACACAGACAGCTCAGCTTACCTGTTACTCAACGAGAGAGT 2686
 QY 2783 CAATGGTACTTGTCTTACGCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2842
 Db 2687 CGGCGGTACTTGTCTTACG-ATTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2745
 QY 2843 CAACCAACAAAGAAATTTTCAATGAAAGATGATGTTGCTTTTGCATTTG 2902
 Db 2746 CAGCCAAACAAAG-AAAAATCCCAAAAGAAAGAGAGAGAGAGAGAGAGAGAGAG 2800
 QY 2903 TGTATAGGCAATCAATATCAGCAAAATGATCTCTTTTAAAGAAATGAGAG 2962
 Db 2801 TATACAGGCAATCAATATGAGAAATCTTTCTTTTAAAGAAATGAGAGAG 2860
 QY 2963 AAAGTGAAGATTTTCCAGAGTGTGAGCCAGGCGCTTAAATTCACAGATTTTAAAG 3022
 Db 2861 AAATAGCAATTTATAGAGAGTGTGAGCCAGGCGCTTAAATTTACAGATTTTAAAG 2920
 QY 3023 AGAAAAACACAGAG--AGAGTACCTCAGGTGTTT 3059
 Db 2921 TGAGAAAAACACAGAAATTAAGCTTACCTCAGGTGTTT 2960
 RESULT 4
 BM561242 999 bp mRNA linear EST 20-FEB-2002
 LOCUS BM561242 6567353 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5739550
 DEFINITION AGENCOURT 6567353 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5739550
 5', mRNA sequence.
 ACCESSION BM561242

VERSION BM561242.1 GI:18806337
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 999)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at:
 http://image.liml.gov
 Plate: LIML12753 row: a column: 23
 High quality sequence stop: 679.
 FEATURES
 source location/Qualifiers
 1..999
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5739550"
 /tissue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_88"
 /note="Organ: small intestine; Vector: pCMV-SPORT6; Site 1: Not1; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
 ORIGIN
 Query Match 28.7%; Score 941.6; DB 12; Length 999;
 Best Local Similarity 98.1%; Pred. No. 7.9e-174;
 Matches 983; Conservative 0; Mismatches 16; Indels 3; Gaps 3;
 QY 1029 CCGGAAAGAACCATCATCTGTGAAGAGGACAGTGTGAGGCTGTGACAGTGTGAGTAA 1088
 Db 1 CCGGAAAGAACCATCATCTGTGAAGAGGACAGTGTGAGGCTGTGACAGTGTGAGTAA 60
 QY 1089 GATTATGAAGAGCTGCGTGAAGGCTTTGAAGATGATGCTGTGTTAACACCACTC 1148
 Db 61 GATTATGAAGAGCTGCGTGAAGGCTTTGAAGATGATGCTGTGTTAACACCACTC 120
 QY 1149 CGGATACTTCTCAGCCTGTACCCCATCAGCAAGTTGGCCGCTTCCGATCATCATC 1208
 Db 121 CGGATACTTCTCAGCCTGTACCCCATCAGCAAGTTGGCCGCTTCCGATCATCATC 180
 QY 1209 TTATCCAGAGGAGATTTGTAATCTTTCATCCCAACGAGGCTGTGAGGCGCATAT 1268
 Db 181 TTATCCAGAGGAGATTTGTAATCTTTCATCCCAACGAGGCTGTGAGGCGCATAT 240
 QY 1269 CGGAAAGAGGGGACACATCAACAGCTGGGAGATTCGCGAGACCTCTATCAAGAT 1328
 Db 241 CGGAAAGAGGGGACACATCAACAGCTGGGAGATTCGCGAGACCTCTATCAAGAT 300
 QY 1329 TGCCTCTGGGAAAGCGCCAGAGCTCAGCAAGAGATGTGATCATACCGGCGCACCG 1388
 Db 301 TGCCTCTGGGAAAGCGCCAGAGCTCAGCAAGAGATGTGATCATACCGGCGCACCG 360
 QY 1389 AGCCCGATTCAAGCCCGCAGAGAGAGATCTTTGGGAAATGAAGAGAAATCTTTAA 1448
 Db 361 AGCCCGATTCAAGCCCGCAGAGAGAGATCTTTGGGAAATGAAGAGAAATCTTTAA 420
 QY 1449 CCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACAGCTGGCG 1508
 Db 421 CCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACAGCTGGCG 480

```
QY 1509 GGTGATTGGCAAAAGGTGGCAAGACCGTGAACGAACTGCAAGAACTTAACCACTGACAGT 1568
Db 481 GGTGATTGGCAAAAGGTGGCAAGACCGTGAACGAACTTAACCACTGACAGT 540
QY 1569 CATGTGCTCTGTGACCAAAAGCCGATGAAATAGAGAAATGATGCTGCAAGATTATTCGG 1628
Db 541 CATGTGCTCTGTGACCAAAAGCCGATGAAATAGAGAAATGATGCTGCAAGATTATTCGG 600
QY 1629 GCACTTCTTGTCTAGCCAGACTGCAAGAGGCAAGATGAGGAAATTTGACAAACAGTGA 1688
Db 601 GCACTTCTTGTCTAGCCAGACTGCAAGAGGCAAGATGAGGAAATTTGACAAACAGTGA 660
QY 1689 GCAGCAGAGAGAGAAATACCTTCAGAGAGTGCCTTCACAGGCGACAAAGTAGGCTCCCA 1748
Db 661 GCAGCAGAGAGAGAAATACCTTCAGAGAGTGCCTTCACAGGCGACAAAGTAGGCTCCCA 720
QY 1749 CAGGCACCAAGCAAAACAAACGATGATATGATGCTTCCACACCTGACAGATGAGCA 1808
Db 721 CAGGCACCAAGCAAAACAAACGATGATATGATGCTTCCACACCTGACAGATGAGCA 780
QY 1809 AACGACCCAGACCGAGATCGGGAGCAAAACAAAGACACTTGAGGAAATGAGAACTTCGG 1868
Db 781 AACGACCCAGACCGAGATCGGGAGCAAAACAAAGACACTTGAGGAAATGAGAACTTCGG 839
QY 1869 AGGCGGCGAGGAGACTCTGCGAGGCGCTTCAGAAACCCGAGGCGCGAGAGGCGGAGAA 1928
Db 840 AGGCGGCGAGGAGACTCTGCGAGGCGCTTCAGAAACCCGAGGCGCGAGAGGCGGAGAA 898
QY 1929 GGTACGACGAGTTTGGCAAGACCAACGAGGCGCGCTTCGCGCGCGCGCGCGCTTCGAG 1988
Db 899 AGTACACCCGAGTTTGGCAAGACCAACGAGGCGCGCTTCGCGCGCGCGCGCGCTTCGAG 958
QY 1989 GCTTACGACCACTTCACATTCACATTCACATTCACATTCACATTCACATTCACATTC 2030
Db 959 GCTTACGACCACTTCACATTCACATTCACATTCACATTCACATTCACATTCACATTC 999
```

```
RESULT 5
LOCUS B0691878
DEFINITION AGENCOURT_8034698 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6207026
5', mRNA sequence.
ACCESSION B0691878
VERSION B0691878.1 GI:21817194
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 874)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2362 row: h column: 03
High quality sequence step: 653.
Location/Qualifiers
1. 874
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6207026"
/issue_type="ductal carcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/clone_id="NIH_MGC_110"
```

FEATURES

source

```
ORIGIN
Query Match 25.7%; Score 842.2; DB 13; Length 874;
Best Local Similarity 99.4%; Pred. No. 2.3e-154;
Matches 856; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

/Note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: This is a NIH MGC Library."

1695 GGACCAAGAAATACCTTCAGAGAGTGCCTTCACAGCGAGCAAGTAGAGCTCCACAGGCA 1754
Db 1 GGACCAAGAAATACCTTCAGAGAGTGCCTTCACAGCGAGCAAGTAGAGCTCCACAGGCA 60
QY 1755 CCAGCAAAACAAACGATGAAATGTAAGCCCTTCACACCTGACAGATGAGACCAAGCA 1814
Db 61 CCAGCAAAACAAACGATGAAATGTAAGCCCTTCACACCTGACAGATGAGACCAAGCA 120
QY 1815 GCCAGCCAGATCGGGAGCAAAACCAAGACCACTGAGAAATGAGAACTTCGGAGGCGG 1874
Db 121 GCCAGCCAGATCGGGAGCAAAACCAAGACCACTGAGAAATGAGAACTTCGGAGGCGG 180
QY 1875 CCAAGGACTTCTCCGAGGCGCTTCAGAAACCCAGGCGCGAGAGGCGGAGAGTCA 1934
Db 181 CCAAGGACTTCTCCGAGGCGCTTCAGAAACCCAGGCGCGAGAGGCGGAGAGTCA 240
QY 1935 CCAAGTTTGGCAAGACCAACGAGGCGCGCTTCGCGCGCGCGCGCGCTTCGAGGCTTCA 1994
Db 241 CCAAGTTTGGCAAGACCAACGAGGCGCGCTTCGCGCGCGCGCGCGCTTCGAGGCTTCA 300
QY 1995 GCCATCCACTTCACATTCACATTCACATTCACATTCACATTCACATTCACATTC 2054
Db 301 GCCATCCACTTCACATTCACATTCACATTCACATTCACATTCACATTCACATTC 360
QY 2055 GTTGAACCTAACATAGGTGAACGTGTTCAAAGCCAGCAAAATGACACCTTTTCTGTG 2114
Db 361 GTTGAACCTAACATAGGTGAACGTGTTCAAAGCCAGCAAAATGACACCTTTTCTGTG 420
QY 2115 GCAAATGCTCTCTGATCATGTGTATATTTGAAGGGGAAATGTTAAGATGTGCG 2174
Db 421 GCAAATGCTCTCTGATCATGTGTATATTTGAAGGGGAAATGTTAAGATGTGCG 480
QY 2175 CTGTGGATTACACAGGAGTGCCTGAGGGGATATATTTTGAAGATATATATCAATA 2234
Db 481 CTGTGGATTACACAGGAGTGCCTGAGGGGATATATTTTGAAGATATATATCAATA 540
QY 2235 CTCAACTAACCTCAATTTTAAATCAATTTTAAATTTTCTTTTAAAGAAAGCA 2294
Db 541 CTCAACTAACCTCAATTTTAAATCAATTTTAAATTTTCTTTTAAAGAAAGCA 600
QY 2295 GGGTTTCTAGACTTTAAAGATTAAGTCTTTGGAGGTCTACGGGTGAGAGAGAGCT 2354
Db 601 GGGTTTCTAGACTTTAAAGATTAAGTCTTTGGAGGTCTACGGGTGAGAGAGAGCT 660
QY 2355 TTGAGGCGCACCGGCAAAATTCACCAAGAGGAAATCTGTTGGAAGGACATCCGCG 2414
Db 661 TTGAGGCGCACCGGCAAAATTCACCAAGAGGAAATCTGTTGGAAGGACATCCGCG 720
QY 2415 AGTTCTGATCACTGTGTATGTCAACAGAGGAGTACCGTCTCTTGAAGAGAACTC 2474
Db 721 AGTTCTGATCACTGTGTATGTCAACAGAGGAGTACCGTCTCTTGAAGAGAACTC 780
QY 2475 TGTCACTCTCATGCTGTCTAGCTATACACCATTTCTTTGCTTACAGGTTTAA 2534
Db 781 TGTCACTCTCATGCTGTCTAGCTATACACCATTTCTTTGCTTACAGGTTTAA 840
QY 2535 A--CTGCTTTTGTGACTAGT 2553
Db 841 AACTGGGTTTTTGTGACTAGT 861
```

RESULT 6	851 bp	mRNA	linear	EST 02-MAY-2002
LOCUS	BQ221568			
DEFINITION	AGNCOURT 7559207 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6046477			
ACCESSION	BQ221568			
VERSION	BQ221568.1			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 851)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-x@mail.nih.gov Tissue Procurement: ATCC/DCTD/DPF cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov Plate: LIML1392 row: f column: 14 High quality sequence stop: 637. Location/Qualifiers 1..851 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6046477" /tissue_type="melanotic melanoma" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH MGC 72" /note="Organ: skin; Vector: PCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."			
FEATURES				
source				
ORIGIN				
Query Match	25.3%;	Score 830;	DB 13;	Length 851;
Best Local Similarity	99.3%;	Pred. No. 5.7e-152;		
Matches	844;	Conservative	0;	Mismatches 5; Indels 1; Gaps 1
QY	47	GGGGGCTCTGGGGAAAGAGACGGATGATGACAAAGCTTTACATCGGAGACCTGAGCCCGC	106	
DB	1	GGGGGCTCTGGGGAAAGAGACGGATGATGACAAAGCTTTACATCGGAGACCTGAGCCCGC	60	
QY	107	CGGTACCGCGACGACCTCCGGAGCTCTTTGGGGACAGAAAGCTGCCCTGCGGGAC	166	
DB	61	CGGTACCGCGACGACCTCCGGAGCTCTTTGGGGACAGAAAGCTGCCCTGCGGGAC	120	
QY	167	AGGTCTGTGTAAGTCGGCTAGCGCTTCGTGACTACCCGACCAAGATGAGCCCATCC	226	
DB	121	AGGTCTGTGTAAGTCGGCTAGCGCTTCGTGACTACCCGACCAAGATGAGCCCATCC	180	
QY	227	GGGCGATGAGACCTCTCTGGGTAAAGTGGATTTGATGGGAAATCATGGAAAGTTGAT	286	
DB	181	GGGCGATGAGACCTCTCTGGGTAAAGTGGATTTGATGGGAAATCATGGAAAGTTGAT	240	
QY	287	ACTCACTCTTAAAAAGCTPAGAGAGAGAAAAATTGATTGAAACATCCCTCTCAACC	346	
DB	241	ACTCACTCTTAAAAAGCTPAGAGAGAGAAAAATTGATTGAAACATCCCTCTCAACC	300	
QY	347	TGCAAGTGGAGAGGTGTTGATGACATTTGGCTCAATATGGGACATGAGAAATGTGAAC	406	
DB	301	TGCAAGTGGAGAGGTGTTGATGACATTTGGCTCAATATGGGACATGAGAAATGTGAAC	360	
QY	407	AAATCAACACAGACAGAAACCGCGCTGTCTACATATGCAACAAGAGAAAG	466	

Db	361	AACTCAACACAGACACAGAAACCGCGTGTCAACGTCACATATGCAACAGAAAGAAAG	420
QY	467	CAAAAATAGCCATGGAGAGCTAAAGCGGGCATCAGTTTGAGAACTACTCCCTTCAAGATT	526
Db	421	CAAAAATAGCCATGGAGAAAGCTAAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATT	480
QY	527	CCATCAATCCCGGATGGAAGAGTGAGTCCCTCTTGCCCTCAGCGAGCCAGCGTGGGG	586
Db	481	CCTACATCCCGGATGGAAGAGTGAGTCCCTCTTGCCCTCAGCGAGCCAGCGTGGGG	540
QY	587	ACCACTCTTCCCGGAGCAAGGCCACGCCCTCTGGGGCACTTCTCAGGCCAGACGATTG	646
Db	541	ACCACTCTTCCCGGAGCAAGGCCACGCCCTCTGGGGCACTTCTCAGGCCAGACGATTG	600
QY	647	ATTTCCCGCTGCCGATTCCTGTCCTCCCAACGATTTGTGTGTCATATCGGAAGAAGAG	706
Db	601	ATTTCCCGCTGCCGATTCCTGTCCTCCCAACGATTTGTGTGTCATATCGGAAGAAGAG	660
QY	707	GCTTGACCATAAAGAAATCATCACTAAGAGAACCCAGTCCCGGGTAGATATTCATAGAAAAG	766
Db	661	GCTTGACCATAAAGAAATCATCACTAAGAGAACCCAGTCCCGGGTAGATATTCATAGAAAAG	720
QY	767	AGAACTGTGAGAGCTGACAGAAAGCTGTACACATTCATGCCACCCACAGAGGGAGCTTCTG	826
Db	721	AGAACTGTGAGAGCTGACAGAAAGCTGTACACATTCATGCCACCCACAGAGGGAGCTTCTG	780
QY	827	AAGCATGCCGATGATCTTGAAATCATGCAAGAAAGGCGAGTAGAGCCAAAC-TAGCC	885
Db	781	AAGCATGCCGATGATCTTGAAATCCGTGACAGAAAGGCGAGTAGAGCCAACTTAGGCC	840
QY	886	GAAGAGATTG 895	
Db	841	GABGAGATTG 850	

```

LOCUS      BX644668                               612 bp      mRNA      linear      EST 04-SEP-2003
DEFINITION DKFZp781G0234.r1.781 (synonym: hlcc4) Homo sapiens cDNA clone
            DKFZp781G0234.5', mRNA sequence.
ACCESSION  BX644668
VERSION    BX644668.1  GI:34479001
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 832)
            Lauber,J., Bahr,A., Mewes,H.W., Well,B., Amid,C., Osanger,A.,
            Pobo,G., Han,M., and Wiemann,S.
            EST (lauber,J., Bahr,A., Mewes,H.W., Well,B., Amid,C., et al.)
            Unpublished (2003)
COMMENT    Contract: MIPS
            MIPS
            Ingolstaedter Landster.1, D-85764 Neuberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
            consortium of the German Genome Project.
            No sl sequence available.
            This clone (DKFZp781G0234) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
            Location/Qualifiers
                1..832
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp781G0234"
                /dev_stage="adult"
                /lab_host="DH10B"

```

/clone_lib="781 (synonym: h1cc4)"
/note="Vector: pSport1_sfi; Site_1: SfiI; Site_2: SfiIb;
cDNA-collection"

ORIGIN

Query Match 24.6%; Score 809; DB 13; Length 832;
Best Local Similarity 99.9%; Pred. No. 7.4e-148;
Matches 820; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 1603 GAGGAAGTATGTCAGAAATTAATCGGAGACTTTCTTGTACGACACTGCAACGCGCAAG 1662
DB 13 GAGGAAGTATGTCAGAAATTAATCGGAGACTTTCTTGTACGACACTGCAACGCGCAAG 72
QY 1663 ATCAGGAAATTTGATCAACAGAGTGAAGCAGCAGGAGAGAGAAATCCCTCAGGAGTCCCC 1722
DB 73 ATCAGGAAATTTGATCAACAGAGTGAAGCAGCAGGAGAGAGAAATCCCTCAGGAGTCCCC 132
QY 1723 TCACAGCGCAGCAAGTGAAGTCTCCACAGGAGCAGCAAAACAAACGATGATGAGCC 1782
DB 133 TCACAGCGCAGCAAGTGAAGTCTCCACAGGAGCAGCAAAACAAACGATGATGAGCC 192
QY 1783 TTCCAAACCTGACAGATGAGACCAACGACGACGACGATGCGAGCAACCAAGA 1842
DB 193 TTCCAAACCTGACAGATGAGACCAACGACGACGACGATGCGAGCAACCAAGA 252
QY 1843 CCATCTGAGGAATGAGAAGTCTGCGGAGGCGCGGAGGAGCTTCCGAGGCGCTGAGAAC 1902
DB 253 CCATCTGAGGAATGAGAAGTCTGCGGAGGCGCGGAGGAGCTTCCGAGGCGCTGAGAAC 312
QY 1903 CCCAGGGGCGCAGAGGAGGCGGAGGAGTCAAGCAGGTTTGCAGAAACCAACGAGCCCG 1962
DB 313 CCCAGGGGCGCAGAGGAGGCGGAGGAGTCAAGCAGGTTTGCAGAAACCAACGAGCCCG 372
QY 1963 CTTCCGCGCGCGCGGCGCTTCTGAGGCTTCAAGCATTCACCTTACCATCTCGGATC 2022
DB 373 CTTCCGCGCGCGCGGCGCTTCTGAGGCTTCAAGCATTCACCTTACCATCTCGGATC 432
QY 2023 TCTCTGAACTCCCAAGCCTATCCCTTTTGTAGTGAATCACTAGTGAAGCTTTCA 2082
DB 433 TCTCTGAACTCCCAAGCCTATCCCTTTTGTAGTGAATCACTAGTGAAGCTTTCA 492
QY 2083 AAGCCAAAGCAAAATGCAACCCCTTTTCTGTGGCAATCGTCTGTACATGTGTACA 2142
DB 493 AAGCCAAAGCAAAATGCAACCCCTTTTCTGTGGCAATCGTCTGTACATGTGTACA 552
QY 2143 TATTAGAAAGGAAAGATGTTAAGATATGTGGCTGTGGTTACAAGGCTGCGACG 2202
DB 553 TATTAGAAAGGAAAGATGTTAAGATATGTGGCTGTGGTTACAAGGCTGCGACG 612
QY 2203 GTAATATATTTAGAAATATATATCAATTAACCTCAACTCCAAATTTTAAATCAAT 2262
DB 613 GTAATATATTTAGAAATATATATCAATTAACCTCAACTCCAAATTTTAAATCAAT 672
QY 2263 ATTAATTTTTTTTTTTTTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTAAAGATAAGT 2322
DB 673 ATTAATTTTTTTTTTTTTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTAAAGATAAGT 752
QY 2323 CTTTGGAGAGTCTCAAGGTGTAGAGAGAGCTTTAGAGCCACCCGCAAAATTTCAACCA 2382
DB 733 CTTTGGAGAGTCTCAAGGTGTAGAGAGAGCTTTAGAGCCACCCGCAAC 791
QY 2383 GAGGGAATCTCGTCGAGAGAGACACTCAAGGAGTCTGGA 2423
DB 792 GAGGGAATCTCGTCGAGAGAGACACTCAAGGAGTCTGGA 832
```

RESULT 8 836 bp mRNA linear EST 02-MAY-2002
LOCUS BQ212500
DEFINITION AGENCODE_7675661 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6095809
ACCESSION BQ212500
VERSION BQ212500.1 GI:20392782

KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 836)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DRP

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNI)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LILNI at:
<http://image.lnl.gov>
Plate: LILNI3368 row: n column: 02
High quality sequence stop: 670.
Location/Qualifiers

FEATURES

1..836
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6095809"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-Sport6; Site_1: NciI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 24.3%; Score 799.2; DB 13; Length 836;
Best Local Similarity 98.7%; Pred. No. 6.2e-146;
Matches 826; Conservative 0; Mismatches 9; Indels 2; Gaps 2;

```
QY 2039 GAGGCTATCCCTTTTGTAGTGAATCACTAGTGAAGTGAAGCTTTCAAGCAAGCAAAATGC 2098
DB 1 GAGGCTATCCCTTTTGTAGTGAATCACTAGTGAAGTGAAGCTTTCAAGCAAGCAAAATGC 60
QY 2099 ACAACCTTTTCTGTGGCAAAATGCTCTGTACATGTGTACATATTTAGAAAGGAGAA 2158
DB 61 ACAACCTTTTCTGTGGCAAAATGCTCTGTACATGTGTACATATTTAGAAAGGAGAA 120
QY 2159 TGTTAAGATATGTGGCTGTGGTTACAGAGGTGCTGAGGCTATATTTTGA 2218
DB 121 TGTTAAGATATGTGGCTGTGGTTACAGAGGTGCTGAGGCTATATTTTGA 180
QY 2219 ATTAATATCAATTAATCACTCACTCAATTTTAAATCAATTTTAAATTTTCT 2278
DB 181 ATTAATATCAATTAATCACTCACTCAATTTTAAATCAATTTTAAATTTTCT 240
QY 2279 TTTTAAAGAAAGCAGGCTTTTCTAGACTTTAAAGATTAAGTCTTTGGAGTCTCAC 2338
DB 241 TTTTAAAGAAAGCAGGCTTTTCTAGACTTTAAAGATTAAGTCTTTGGAGTCTCAC 300
QY 2339 GGTGTAGAGAGAGCTTTAGAGCAGCCGCAAAATTTACCCAGAGGAAATCTGTGC 2398
DB 301 GGTGTAGAGAGAGCTTTAGAGCAGCCGCAAAATTTACCCAGAGGAAATCTGTGC 360
QY 2399 GAAGCACTCAAGCAGTCTGATCACTGTGTATGTCAACAGAGGAAATCTGTGC 2458
DB 361 GAAGCACTCAAGCAGTCTGATCACTGTGTATGTCAACAGAGGAAATCTGTGC 420
QY 2459 CTTGAAGAGAAATCTGTCACTCTCATGCTGTCTAGCTATACACCATTTCTCTT 2518
DB 421 CTTGAAGAGAAATCTGTCACTCTCATGCTGTCTAGCTATACACCATTTCTCTT 480
QY 2519 GCTCAAGGTTTAACTGTTTTTTTGTATCTGTATATATTTCTCTCTCTCTG 2578
```

Db 481 GCTTCACAGGTTTAACTGGTTTTCATCTATGATATATCTCTGCTCTCTG 540
QY 2579 TTTATCTCTCCCTCCCTCCCTCCCTCTTCTTCATCTCATCTTTTGAATTTCTC 2638
Db 541 TTTATCTCTCCCTCCCTCCCTCCCTCTTCTTCATCTCATCTTTTGAATTTCTC 600
QY 2639 ATCCCTCATCTCAATCCCGTATCTAGACACCCCGCCCGCCGAGCAAGCTGCTCT 2698
Db 601 ATCCCTCATCTCAATCCCGTATCTAGACACCCCGCCCGCCGAGCAAGCTGCTCT 659
QY 2699 GAGTATCATCTACACAAAGGAAACAAAGGAAACACAAACAGCTCACTTAC 2758
Db 660 GAGTATCATCTACACAAAGGAAACAAAGGAAACACAAACAGCTCACTTAC 719
QY 2759 TTGGTTACTCAAAAGAACAGTCAATGTATCTTGTCTTA-GGTTTGGAGAGGAA 2817
Db 720 TTGGTTACTCAAAAGAACAGTCAATGTATCTTGTCTTANCGTTTGGAGAGGAA 779
QY 2818 ACAGGAACCCACCAACCAACCAATCAACCAACCAAAATTCACATGAAA 2874
Db 780 ACAGGAACCCACCAACCAACCAATCAACCAACCAAAATTCACATGAAA 836

RESULT 9
BM806132 1089 bp mRNA linear EST 05-MAR-2002
LOCUS BM806132
DEFINITION AGENCOURT 6553922 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5556552
5', mRNA sequence.
ACCESSION BM806132
VERSION BM806132.1 GI:19122955
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1089)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12376 row: c column: 13
High quality sequence stop: 688.

FEATURES
Source
1..1089
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:555652"
/issue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_host="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

ORIGIN
Query Match 24.3%; Score 797; DB 12; Length 1089;
Best Local Similarity 94.3%; Pred. No. 1.6e-145;
Matches 905; Conservative 0; Mismatches 41; Indels 14; Gaps 7;

QY 201 CTAACCCGACGCAAGTGGCCCATCGCGCATCGAGACCCCTCGGGTAAATGGAAT 260
Db 1 CTAACCCGACGCAAGTGGCCCATCGCGCATCGAGACCCCTCGGGTAAATGGAAT 60
QY 261 GCATGGAAATCATGGAATGATTACTCACTCTAAAGCTAAGAGCAGGAAAT 320

Db 61 GCATGGAAATCATGGAATGATTACTCACTCTAAAGCTAAGAGCAGGAAAT 120
QY 321 TCGAATTGGAACATCCCTCCCTCACTGAGTGGAGAGTGTGGATCTTTGGCTCA 380
Db 121 TCGAATTGGAACATCCCTCCCTCACTGAGTGGAGAGTGTGGATCTTTGGCTCA 180
QY 381 ATATGGACAGTGGAGATGTGGAACAATGCAACAGACACAGAAACCGCGTGTCA 440
Db 181 ATATGGACAGTGGAGATGTGGAACAATGCAACAGACACAGAAACCGCGTGTCA 240
QY 441 CGTCACTATGCAACAGAAAGAAATAGCCATGGAAGAGTGAAGCGGACATCA 500
Db 241 CGTCACTATGCAACAGAAAGAAATAGCCATGGAAGAGTGAAGCGGACATCA 300
QY 501 GTTGGAGAACTACTCTTCAAGATTTCTCATCTCCGATGAAGAGTGAAGCTCCCTTC 560
Db 301 GTTGGAGAACTACTCTTCAAGATTTCTCATCTCCGATGAAGAGTGAAGCTCCCTTC 360
QY 561 GCCCCTCAGGAGCCGAGCGTGGGACCACTCTTCCCGGAGCAGGCGCCCTGG 620
Db 361 GCCCCTCAGGAGCCGAGCGTGGGACCACTCTTCCCGGAGCAGGCGCCCTGG 420
QY 621 GGGCACTTCTAAGCCAGACGATTTGATTCCTGCGGATCCGTCGCCACCAAGTT 680
Db 421 GGGCACTTCTAAGCCAGACGATTTGATTCCTGCGGATCCGTCGCCACCAAGTT 480
QY 681 TGTGGTGCATCATCGGAAGAGAGGCTTGACATTAAGAACATCACTAAGCAGACCA 740
Db 481 TGTGGTGCATCATCGGAAGAGAGGCTTGACATTAAGAACATCACTAAGCAGACCA 540
QY 741 GTCCCGGTGATATCATTAAGAAAGAGAACTCTGAGCTGAGAGAGGCTGTACCAT 800
Db 541 GTCCCGGTGATATCATTAAGAAAGAGAACTCTGAGCTGAGAGAGGCTGTACCAT 600
QY 801 CCATGCGCCCGCAGAGGGGACCTTGAGAGCATGCGCATGATCTTGAATATGACAGA 860
Db 601 CCATGCGCCCGCAGAGGGGACCTTGAGAGCATGCGCATGATCTTGAATATGACAGA 660
QY 861 AGAGGCATGAGACCAACTAGCCGAGAGATTCCTGTGAATCTTGGCAGACATGG 920
Db 661 AGAGGCATGAGACCAACTAGCCGAGAGATTCCTGTGAATCTTGGCAGACATGG 720
QY 921 --CTTGTGGAGAGCTGATTTGAAAGAGAGGAGCAAAATTTGAAGAAATGGAATGTA 978
Db 721 GCTTGTTGGAGAGCTGATTTGAAAGAGAGGAGCAAAATTTGAAGAAATGGAATGTA 780
QY 979 ACA-GGACCAAGATTAACATCTCTT-GCAGGATTTGAGCATATCAACCCGG-AA 1035
Db 781 ACA-GGACCAAGATTAACATCTCTT-GCAGGATTTGAGCATATCAACCCGGAA 840
QY 1036 AGAACCATCACTG-TGAAGGGCAGAG-TTGAAGGCTGTGCAAGTGAATGAGATTA 1093
Db 841 AGAACCATCACTGTTGAAGGGCAGAGTTGAGCCCTGTGCCATGCTGAAGATGAGAT 900
QY 1094 TGAAGAGCTGC-----GTAGAGCTTTGAAATGATATCTGCTTTAAACACCA 1146
Db 901 TATGGAAGAACTGTGCGCGAGAGGCCCTTTGAAATGAAATGCTGAGCTTTAAACCA 960

RESULT 10
BU156245 894 bp mRNA linear EST 03-SEP-2002
LOCUS BU156245
DEFINITION AGENCOURT 7965128 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6162660
5', mRNA sequence.
ACCESSION BU156245
VERSION BU156245.1 GI:22669777
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 894)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov

Tissue Procurement: ATCC/DCTD/DP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>

Plate: L1AM1316 row: 0 column: 13
High quality sequence stop: 666.

FEATURES
source location/Qualifiers

1..894

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6162660"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_72"

/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 23.1%; Score 759.8; DB 13; Length 894;
Best Local Similarity 95.5%; Pred. No. 3.2e-138;
Matches 836; Conservative 0; Mismatches 32; Indels 7; Gaps 5;

```

QY 1973 CCCAGGGCTTGGAGGCTTCAGCATCCATTCATCCATCGGATCTCTCTGAAC 2032
DB 20 CCTCTCTACCTGAGGCTTCAGCATCCATTCATCCATCGGATCTCTCTGAAC 79
QY 2033 TCCACAGAGCTATCCCTTTTGAAGTGAACATAGTGAACGTTTCAAGCCAGCA 2092
DB 80 TCCACAGAGCTATCCCTTTTGAAGTGAACATAGTGAACGTTTCAAGCCAGCA 139
QY 2093 AAATGCAACCCCTTTTGTGGCAATGCTCTGTACATGTGTACATATTAGAAAG 2152
DB 140 AAATGCAACCCCTTTTGTGGCAATGCTCTGTACATGTGTACATATTAGAAAG 139
QY 2153 GGAAGATGTAATATGTGGCTGTGGGTTACACAGGGTGCTGCAAGCGGTAAATATT 2212
DB 200 GGAAGATGTAATATGTGGCTGTGGGTTACACAGGGTGCTGCAAGCGGTAAATATT 259
QY 2213 TTGAATATATATATCAATATCACTCACTCACTCACTTTTATCATATTAATTTT 2272
DB 260 TTGAATATATATATCAATATCACTCACTCACTCACTTTTATCATATTAATTTT 319
QY 2273 TTTTCTTTTAAAGAGAAAGAGGCTTTTCTAGACTTTTAAAGATTAAGCTTTGGAGG 2332
DB 320 TTTTCTTTTAAAGAGAAAGAGGCTTTTCTAGACTTTTAAAGATTAAGCTTTGGAGG 379
QY 2333 TGTCAAGGTGTAGAGGAGCTTTTGAAGCCACCCGCAAAATATCCACCGAGGAAATC 2392
DB 380 TGTCAAGGTGTAGAGGAGCTTTTGAAGCCACCCGCAAAATATCCACCGAGGAAATC 439
QY 2393 TGTGCGAAGACACTCACGCGAGTCTGGATCACTGTGTATGTCAACAGAAAGGATTC 2452
DB 440 TGTGCGAAGACACTCACGCGAGTCTGGATCACTGTGTATGTCAACAGAAAGGATTC 499
QY 2453 CGTCTCTTGAAGAGAAACTCTGTCACTCTCATGCTGTAGCTATACCCATTT 2512
DB 500 CGTCTCTTGAAGAGAAACTCTGTCACTCTCATGCTGTAGCTATACCCATTT 559
QY 2513 CTCTTGTCTTCAAGGTTTAAATCTGTTTGTGATCTGCTATTAATTTCTGTCTC 2572
DB 560 CTCTTGTCTTCAAGGTTTAAATCTGTTTGTGATCTGCTATTAATTTCTGTCTC 619

```

```

QY 2573 TCTCGTTTATCTCTCCCTCCCTCCCTCCCTCCCTCTTCTCTCATCTCA--TTCTTTTGA 2631
DB 620 TCTCTGTATCTCTCCCTCCCTCCCTCCCTCCCTCTTCTCTCATCTCTCTTTTGA 679
QY 2632 TTTCTCATCCCTCATCTCAATCCCGTATCTAGCAACCCCGCCCCCGGAGCAAGCA 2691
DB 680 TTTCTCATCCCTCATCTCAATCCCGTATCTAGCA--CCCCCCCCCGGAGCAAGCA 737
QY 2692 GTGCTGTAGTATCATATCACAAAGAAACAAAGGCAACACACAAACAGCTTCA 2751
DB 738 GTGCTGTAGTATCATATCCCCCAAAAGGAGCAAAAGGAAACACACACAGCTTCA 797
QY 2752 C-TTACACTTGTTACTTCAAAAG-AAACAAGTCAATGTACTGTCTAGCGTTTGG- 2808
DB 798 CTTTACACTTGTTACTTCAAAAGAAACAAAGGAGGTTAATTGCTTACCGTTTGG 857
QY 2809 -AAGAGAAAAAGAGAACCCACCAACCAACCAAT 2842
DB 858 AAGAGAAAAAGAGAACCCACCAACCAACCAAT 892

```

RESULT 11
BG575889 781 bp mRNA linear EST 10-APR-2001
LOCUS BG575889
DEFINITION 602598315F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4707243 5', mRNA sequence.
ACCESSION BG575889
VERSION BG575889.1 GI:13583542
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Bkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DP
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: L1AM10576 row: m column: 04
High quality sequence stop: 773.

FEATURES

source

1..781

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4707243"

/tissue_type="mammary adenocarcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_87"

/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC library."

ORIGIN

Query Match 22.9%; Score 752.6; DB 12; Length 781;
Best Local Similarity 98.6%; Pred. No. 8.3e-137;
Matches 770; Conservative 0; Mismatches 9; Indels 2; Gaps 1;

```

QY 1692 GCAGAGCAGAAATACCTCAGGAGTGCCTCACAGCCAGCAAGTGAAGCTCCACAG 1751
DB 1 GCAGAGCAGAAATACCTCAGGAGTGCCTCACAGCCAGCAAGTGAAGCTCCACAG 60
QY 1752 GCAGCAGAAACAGATGATGTAGCCCTTCCAACTGACAGAAATGAGACCAAC 1811

```


Db	688	AGAAAGCCCTGTACACATCCATCCATCCAGCCAGAGGGAGACTTCTGAGACATGCCGATATTC	747
QY	845	TTGAAATCATTCGA-GAAAGAGCGCATGTAGACCAACTAG-CCGAGAGATT-CCCTCTGA	901
Db	748	TTGAAATCATTCGAGGAGAGGAGCGATGAGACCAACTAGCCGAGAGATTCCTCTGA	807
QY	902	AAATCTGG	910
Db	808	AAATCTGG	816

RESULT 13	CF593505	824 bp	mRNA	linear	EST 26-SEP-2001
LOCUS	AGENCOURT_15622249 NIH MGC_147 Homo sapiens cDNA clone				
DEFINITION	IMAGE:30531076 5', mRNA sequence.				
ACCESSION	CF593505				
VERSION	CF593505.1	GI:36347108			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	NIH-MGC http://mgi.nci.nih.gov/				
JOURNAL	1 (bases 1 to 824)				
COMMENT	Unpublished (1999)				
	Contact: Daniela S. Gerhard, Ph.D.				
	Office of Cancer Genomics				
	National Cancer Institute / NIH				
	Bldg. 31 Rm10A07 Bethesda, MD 20892				
	Email: cgapbs-rt@mail.nih.gov				
	Tissue Procurement: Dr. Stefan Hanson				
	cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help				
	and advice from Piero Carninci (RIKEN)				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNL at:				
	http://image.lnl.gov				
	Plate: NDM621 row: e column: 05				
	High quality sequence stop: 662.				
	Location/Qualifiers				
	1. 824				
	/organism="Homo sapiens"				
	/mol type="mRNA"				
	/db xref="taxon:9606"				
	/clone="IMAGE:30531076"				
	/tissue type="Human Placenta"				
	/lab host="DH10B Tona"				
	/clone.lib="NIH MGC 147"				
	/note="Organ: placenta; Vector: pBluescript; Site_1:				
	all-XhoI; Site_2: BamH; Oligo-dT primed using primer				
	5'-TTTTTTTTTTTTTTT-3', size-selected for average				
	insert size 2.3 kb and normalized to ROT 5. This is a				
	primary library enriched for full-length clones and				
	constructed using the Cap-trapper method (Carninci, in				
	preparation). Library constructed by M. Brownstein				
	(NIH/NHGRI, National Institutes of Health). Note: This is				
	a NIH_MGC library."				

FEATURES	SOURCE
Query Match	22.8%; Score 749.8; DB 14; Length 824;
Best Local Similarity	99.4%; Pred. No. 2.9e-116;
Matches	773; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

ORIGIN
1 GGCAAGCGAGAGAGCGAGAGCGCCGAGTACCGGAGCGAGAGCGCGGCTCTCGGG 60
38 GGCAGCGAGAGAGCGAGAGCGCCGAGTACCGGAGCGAGAGCGCGGCTCTCGGG 97
61 AAGAGAGGAGTGAAGAAAGCTTAACTCGGGAACCTGAGCCCGCGCTACCGCGAGC 120
98 AAGAGAGGAGTGAAGAAAGCTTAACTCGGGAACCTGAGCCCGCGCTACCGCGAGC 157

QY	121	GACCTCCGGCACTCTTTGGGAGCAAGAAAGCTGCCCTTGGCGGACAGCTCCGCTGAAG	180
Db	158	GACCTCCGGCACTCTTTGGGAGCAAGAAAGCTGCCCTTGGCGGACAGCTCCGCTGAAG	217
QY	181	TCGGGCTACGCCCTTGTGACTACCCCGACAGAACTGGGCACTCCGCACTGAGACC	240
Db	218	TCGGGCTACGCCCTTGTGACTACCCCGACAGAACTGGGCACTCCGCACTGAGACC	277
QY	241	CTCTGGGCTAAAGTGGAAATTGCAATGGGAAAATCATGAAATTGATTACTGCTTAA	300
Db	278	CTCTGGGCTAAAGTGGAAATTGCAATGGGAAAATCATGAAATTGATTACTGCTTAA	337
QY	301	AAGCTAAGGAGCAGGAAAATTCAGATTGCAATCCCTCTCACTGCAAGTGGAGG	360
Db	338	AAGCTAAGGAGCAGGAAAATTCAGATTGCAATCCCTCTCACTGCAAGTGGAGG	397
QY	361	TTGATGAGACTTTTGGCTCAATATGAGGACAGTGAAGATGTGGAAACAAGTCAACAGAC	420
Db	398	TTGATGAGACTTTTGGCTCAATATGAGGACAGTGAAGATGTGGAAACAAGTCAACAGAC	457
QY	421	AAGAAACCGCGCTGTGAAAGTCAAGTCAATATGCAACAAGAAAGCAAAAATAGCCATG	480
Db	458	AAGAAACCGCGCTGTGCAACGTCAATATGCAACAAGAAAGCAAAAATAGCCATG	517
QY	481	GAGAAAGTAAAGCGGCACTAGTTGAGAACTACTGCTTCAAGATTTCATCATCCCGAT	540
Db	518	GAGAAAGTAAAGCGGCACTAGTTGAGAACTACTGCTTCAAGATTTCATCATCCCGAT	577
QY	541	GAGAAAGTGAAGTCCCTTTCGCCCCCTGACGAGCCAGCGTGGGACCACTTTCCCGG	600
Db	578	GAGAAAGTGAAGTCCCTTTCGCCCCCTGACGAGCCAGCGTGGGACCACTTTCCCGG	637
QY	601	GAGCAAGGCGACGCGCCCTGGGGGCACTTCTAGGCGACACAGATGATTTCCCGCTGGG	660
Db	638	GAGCAAGGCGACGCGCCCTGGGGGCACTTCTAGGCGACACAGATGATTTCCCGCTGGG	697
QY	661	ATCCTGTGTCCCAACCCAGTTTGTGTGCCATCATCGGAAAGAGGGCTTGACATTAAG	720
Db	698	ATCCTGTGTCCCAACCCAGTTTGTGTGCCATCATCGGAAAGAGGGCTTGACATTAAG	757
QY	721	AACATCACTTAAGCAGACCCAGTCCG-GGGTGATATTCGATA-GAAAAAGAACTCTGG	776
Db	758	AACATCACTTAAGCAGACCCAGTCCGCGGGTGTGATATTCCTTGAAGAAAAAGAACTCTGG	815

RESULT 14

Seq748346

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homosapiens

Homosapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 889)

NIH-MGC htcp://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-remail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Plate: LICM1677 row: c column: 02

High quality sequence stop: 773.

FEATURES

source

Location/Qualifiers

1..889
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:484253"
 /tissue_type="normal pigmented retinal epithelium"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_43"
 /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library." "

ORIGIN

Query Match 22.7%; Score 746; DB 12; Length 889;
 Best Local Similarity 95.5%; Pred. No. 1,6e-135;
 Matches 802; Conservative 0; Mismatches 30; Indels 8; Gaps 3;

```

72  GATGAACAAGCTTTTACATCGGGAACTGAGCCCGCGTCAACCGCCGACGACCTCCGGCA 131
    |||||
2  GATGAACAAGCTTTTACATCGGGAACTGAGCCCGCGTCAACCGCCGACGACCTCCGGCA 61
132 GCTCTTTGGGAGACAGAAAGCTCCCTCGGAGACAGGTCTCTGAGAGTCCGGCTACGC 121
    |||||
62 GCTCTTTGGGAGACAGAAAGCTCCCTCGGAGACAGGTCTCTGAGAGTCCGGCTACGC 121
132 CTTCTGTGACCTACCCGACCAAGACTGGGCGCATCCGCGCATCGAGACCTCCGGCTAA 251
    |||||
122 CTTCTGTGACCTACCCGACCAAGACTGGGCGCATCCGCGCATCGAGACCTCCGGCTAA 181
252 AGTGAATTGCATGGGAAATCATGGAAAGTTGATTAAGTCTCTAAAGCTAAAGAG 311
132 AGTGAATTGCATGGGAAATCATGGAAATCATGGAAATGATTAAGTCTCTAAAGCTAAAGAG 241
    |||||
312 CAGGAAATTCAGATTGGAACATCCCTCTCACTGACGTGGAGAGTGTGGATGACT 371
242 CAGGAAATTCAGATTGGAACATCCCTCTCACTGACGTGGAGAGTGTGGATGACT 301
372 TTTGGCTCAATTTGGAGAGTGGAGAAATGGAGAACAGTCAACAGACAGACAGAAACGCC 431
    |||||
302 TTTGGCTCAATTTGGAGAGTGGAGAAATGGAGAACAGTCAACAGACAGACAGAAACGCC 361
432 CATTGTCAACGTCAATATGCAACAGAGAGAGCAAAATATAGCCATGAGAGCTTAAG 491
362 CATTGTCAACGTCAACATATGCAACAGAGAGAGCAAAATATAGCCATGAGAGCTTAAG 421
492 CGGCGCATCATGTTAGAACTACTCTCTTCAAGATTTCTTACATCCCGATGAGAGGTGAG 551
    |||||
422 CGGCGCATCATGTTAGAACTACTCTCTTCAAGATTTCTTACATCCCGATGAGAGGTGAG 481
552 CTCCCTTGGCCCTCAGAGAGCCAGAGGTGGGAGCAACTCTTCCCGGAGAGAGGCA 611
    |||||
482 CTCCCTTGGCCCTCAGAGAGCCAGAGGTGGGAGCAACTCTTCCCGGAGAGAGGCA 541
612 CGCCCTCGGGGGCACTTCTCAGGCGCAGACAGATTTGATTTCCCGTCCGATCTCTGTC 671
    |||||
542 CGCCCTCGGGGGCACTTCTCAGGCGCAGACAGATTTGATTTCCCGTCCGATCTCTGTC 601
672 CACCGATTTGTTGTGTCATCATCGAAAAGAGGGCTTGAACATTAAGAACTCACTAA 731
    |||||
732 CACCGATTTGTTGTGTCATCATCGAAAAGAGGGCTTGAACATTAAGAACTCACTAA 661
    |||||
602 CACCGATTTGTTGTGTCATCATCGAAAAGAGGGCTTGAACATTAAGAACTCACTAA 661
732 GAGAGACCGATCCCGGGTATGATTCATGAGAAAGAACTCGAGCTGAGAGAGCC 791
    |||||
662 GAGAGACCGATCCCGGGTATGATTCATGAGAAAGAACTCGAGCTGAGAGAGAG-C 720
792 TGTCAACATTCATCCACCCAGAGGGGACTTCTGAA--GCATGCGCATGATCTTGA 849
    |||||

```

Db 721 TGTCAACATTCATCCACCCAGAGGGGACTTCTTAAGCATTTGCCGATGATTTCTGTA 780
 Qy 850 ATCATGCAAGAAAGAGC-----AGATGACCAAACTAGCCGAGAGATTCCTCTGAAA 904
 Db 781 AATCTGGACACAGACCGCATTTTGAACCAAAATAGCCGAGAAATGCTCTGTTAA 840

RESULT 15
 BU601634 762 bp mRNA linear EST 20-SEP-2002
 LOCUS AGNCOURT 10049220 NIH_MGC_142 Homo sapiens cDNA clone
 DEFINITION IMAGE:6495584 5', mRNA sequence.
 ACCESSION BU601634
 VERSION BU601634.1 GI:23253393
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 762)
 NIH-MGC http://mgi.nhl.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: NCI
 CDNA Library Preparation: Michael Brownstein Laboratory
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/MLN at:
 http://image.llnl.gov
 Plate: LLCM2673 row: k column: 09
 High quality sequence stop: 591.

FEATURES

source

Location/Qualifiers

1..762
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6495584"
 /tissue_type="mixed (pool of 40 RNAs)"
 /lab_host="DH10B (T1-phage-resistant)"
 /clone_lib="NIH_MGC_142"
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccatnagcc); Site 2: SfiI (ggcgccctggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:
 5'-AAGCAGTGGTATCAAGAGAGTGGGCGCATTAAGCGCGG-3' and
 5'-ATTCTAGAGCCGAGCGCGGCCACAGG-dt(30)NN-3'. Pull-length enriched library was constructed using the Clontech Creator SMART Kit and size-selected to contain the >0.5 kb size fraction (other fractions present in NIH_MGC 141). Library created in the laboratory of M. Brownstein (NIH, NIH). Note: this is a NIH_MGC library." "

ORIGIN

Query Match 21.9%; Score 718.8; DB 13; Length 762;
 Best Local Similarity 99.6%; Pred. No. 3.5e-130;
 Matches 731; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

Qy 1 GCGAGCGAGAGAGCGCGCGGTAACCGGCGCGGGGAGCGCGGGGCTCTGGGG 60
    |||||
Db 7 GCGAGCGAGAGAGCGCGCGGTAACCGGCGCGGGGAGCGCGGGGCTCTGGGG 66
    |||||
Qy 61 AAGAGCGAGTGAATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGTCAACCGCGAC 120
    |||||
Db 67 AAGAGCGAGTGAATGAACAAGCTTTACATCGGGAACCTGAGCCCGGTGCAACCGCGAC 126
    |||||

```

QY 121 GACCTCCGGCAGCTCTTTGGGAGACGAAAGCTGCCCTTGCGGAGACAGTCTGTGAAG 180
 |||||
 Db 127 GACCTCCGGCAGCTCTTTGGGAGACGAAAGCTGCCCTTGCGGAGACAGTCTGTGAAG 186
 |||||
 QY 181 TCCGGCTACGCTTCTGTGACTACCCGACCAAGACTGGGCCATCCGGCCATCGAGACC 240
 |||||
 Db 187 TCCGGCTACGCTTCTGTGACTACCCGACCAAGACTGGGCCATCCGGCCATCGAGACC 246
 |||||
 QY 241 CTCTCGGGTAAAGTGAATGCAATGGGAAATCAATGAAAGTTGATTAATCTAGTCTTAA 300
 |||||
 Db 247 CTCTCGGGTAAAGTGAATGCAATGGGAAATCAATGAAAGTTGATTAATCTAGTCTTAA 306
 |||||
 QY 301 AAGCTAAGAGACAGAAATTCAGATTCGAAACATCCCTCTCACTGCACTGGAGAGTG 360
 |||||
 Db 307 AAGCTAAGAGACAGAAATTCAGATTCGAAACATCCCTCTCTCACTGCACTGGAGAGTG 366
 |||||
 QY 361 TTGATGGAATTTTGGCTCAATATGAGACAGTGGAGATGTGAACAAGTCAACACAGAC 420
 |||||
 Db 367 TTGATGGAATTTTGGCTCAATATGAGACAGTGGAGATGTGAACAAGTCAACACAGAC 426
 |||||
 QY 421 ACAGAAACCGCCGTGTCACGTGCATATGCAACAAGAGAGCAAAATATGCCATG 480
 |||||
 Db 427 ACAGAAACCGCCGTGTCACGTGCATATGCAACAAGAGAGCAAAATATGCCATG 486
 |||||
 QY 481 GAGAGCTAAGCGGGCATCAAGTTGAGAACTACTCTTCAAGATTTCTAATCCCGGAT 540
 |||||
 Db 487 GAGAGCTAAGCGGGCATCAAGTTGAGAACTACTCTTCAAGATTTCTAATCCCGGAT 546
 |||||
 QY 541 GAAGAGGTGAGCTCCCTTGCGCCCTCAGCGAGCCAGCGTGGGAGCAACTTCCCGG 600
 |||||
 Db 547 GAAGAGGTGAGCTCCCTTGCGCCCTCAGCGAGCCAGCGTGGGAGCAACTTCCCGG 606
 |||||
 QY 601 GAGCAAGGCCAGC-CGCCGTGGGGCAATTCTCAGGCCAGACAGATTGATTTCCCGCTCG 659
 |||||
 Db 607 GAGCAAGGCCAGC-CGCCGTGGGGCAATTCTCAGGCCAGACAGATTGATTTCCCGCTCG 666
 |||||
 QY 660 GATCTGTGTCCCAACCAAGTTGTGTGTGCATCATCGAAAGAGAGGCTTGAACATAA 719
 |||||
 Db 667 GATCTGTGTCCCAACCAAGTTGTGTGTGCATCATCGAAAGAGAGGCTTGAACATAA 726
 |||||
 QY 720 GAACATCATTAAGC 733
 |||||
 Db 727 GAACATCATTAAGC 740
 |||||

Search completed: July 24, 2004, 16:42:45
 Job time : 8228.67 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:51:25 ; Search time 144.09 Seconds

(without alignments)
12875.302 Million cell updates/sec

Title: US-09-270-437D-8
Perfect score: 5886
Sequence: 1 ggcagcgagagcgagcgagga.....aaccttgaatgttattt 3283

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p_model -DEV=x1p
-O=/cgn2_1/USPTO.spool.p/US09270437/runat.23072004.095738.27403/app.query.fasta_1.5582
-DB=A_Geneseq.29Jan04 -QPMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.ccd
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437.QCGN_1.1_185.qrunat.23072004.095738.27403 -NCPD=6 -ICPD=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq.29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2949	50.1	594	4	ABG06795
2	2838	48.2	556	5	ABG96346
3	2836.5	48.2	614	4	ABG06794
4	2808	47.7	620	6	AAU16163
5	2808	47.7	620	6	ABU55232
6	2584.5	43.9	555	6	ABU8799
7	1934	32.9	577	2	AAU16164
8	1875	31.9	579	3	AAU16165
9	1875	31.9	579	5	ABU75053
10	1875	31.9	579	5	ABU74997

11	1875	31.9	579	5	ABU75054	ABU75054 Human lun
12	1875	31.9	579	5	ABU61917	ABU61917 Human lun
13	1875	31.9	579	5	ABU61974	ABU61974 Human lun
14	1875	31.9	579	5	ABU61973	ABU61973 Human lun
15	1875	31.9	579	7	ADA28536	ADA28536 Recombina
16	1875	31.9	579	7	ADA28539	ADA28539 Recombina
17	1875	31.9	579	7	ADA28438	ADA28438 Human lun
18	1870	31.8	586	5	ABU75048	ABU75048 Human lun
19	1870	31.8	586	5	ABU61968	ABU61968 Human lun
20	1870	31.8	586	7	ADA28517	ADA28517 Recombina
21	1868	31.7	579	3	AAU13328	AAU13328 Human lun
22	1868	31.7	579	5	ABU74960	ABU74960 Human lun
23	1868	31.7	579	5	ABU61880	ABU61880 Human lun
24	1868	31.7	579	7	ADA28266	ADA28266 Human lun
25	1868	31.7	579	7	ADD14066	ADD14066 Human lun
26	1868	31.7	579	7	ADE53471	ADE53471 Human lun
27	1798	30.5	619	4	ABG21963	ABG21963 Novel hum
28	1725	29.3	583	4	ABG12592	ABG12592 Novel hum
29	1075.5	18.3	319	4	AAU93826	AAU93826 Human pol
30	872	14.8	187	6	AAU23971	AAU23971 Human IGF
31	855	14.5	558	4	ABU58367	ABU58367 Drosophila
32	812.5	13.8	261	4	AAU16161	AAU16161 Human nov
33	812.5	13.8	261	6	ABU55230	ABU55230 Human nov
34	803	13.6	171	6	AAU16166	AAU16166 Human nov
35	803	13.6	171	6	ABU55235	ABU55235 Human nov
36	782.5	13.3	250	4	AAU16579	AAU16579 Human nov
37	782.5	13.3	250	6	ABU55648	ABU55648 Human nov
38	750	12.7	171	4	AAU16583	AAU16583 Human nov
39	750	12.7	171	6	ABU55652	ABU55652 Human nov
40	676.5	11.5	192	4	AAU16164	AAU16164 Human nov
41	676.5	11.5	192	6	ABU55233	ABU55233 Human nov
42	501	8.5	97	3	AAU03261	AAU03261 Human sec
43	473.5	8.0	266	4	ABG12593	ABG12593 Novel hum
44	469.5	8.0	209	4	ABG21961	ABG21961 Novel hum
45	458	7.8	148	4	ABG21962	ABG21962 Novel hum

ALIGNMENTS

RESULT 1
ABG06795
ID ABG06795 standard; protein; 594 AA.

AC ABG06795;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6786.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Dmanac RT, Liu C, Tang YT;
XX
DR WPI, 2001-639362/73.
DR N-PSDB; AAS70982.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

OY	1624	ATCGGAGCACTTGTTCGTAAGCCACACTGCACAGCGCAGAATCAGGGAAATTGTACAACAG	1683
Dd	557	IleGlyHisSpePheAlaSerGlnThrValacinarGlySileArgGluIleValGlnGln	576
OY	1684	GTTAAGCAGCAGCAGCAGCAGCAAAATACCCTCAGAGGAGTGCGCTCACAGGCCGACGACAG	1737
Dd	577	ValtyrGlnGlnGlnGlnLysTyrrProGlnGlnyValAlaSerGlnArgSerLys	594
RESULT 2			
ID	ABG96346		
XX	ABG96346 standard; protein; 556 AA.		
AC	ABG96346;		
DT	11-DEC-2002 (first entry)		
DE	Human ovarian cancer marker M452.		
KW	Human; ovarian cancer; marker; cancer; familial history; brain disorder;		
KW	central nervous system disorder; bacterial meningitis; viral meningitis;		
KW	Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;		
KW	brain herniation; inflammation; encephalitis; testicular disorder;		
KW	nontuberculous granulomatous orchitis; connective tissue disorder;		
KW	heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;		
KX	histological type; carcinogenic; ovarian cancer marker.		
OS	Homo sapiens.		
NN	WO200271928-A2.		
FN	19-SEP-2002.		
XX	14-MAR-2002; 2002WO-US007826.		
PR	14-MAR-2001; 2001US-0276025P.		
PR	10-AUG-2001; 2001US-0311732P.		
PR	19-SEP-2001; 2001US-0323580P.		
PR	26-SEP-2001; 2001US-0324967P.		
PR	26-SEP-2001; 2001US-0325102P.		
PR	26-SEP-2001; 2001US-0325149P.		
PA	(MILL-) MILLENNIUM PHARM INC.		
PI	Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;		
PI	Meyers RE, Morrissett MP, Olandt RJ, Sen A, Vieby PO, Mills GB;		
PI	Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;		
XX	WPI; 2002-723277/78.		
DR	N-PDB; ABST6442.		
PT	Assessing whether a patient is afflicted with ovarian cancer, useful in		
PT	assessing the stage or progression of the disease, comprises comparing		
PT	the expression level of a cancer marker in a sample from a patient and		
PT	from a non cancer patient.		
PS	Disclosure; Page 263-264; 481pp; English.		
XX	The present invention relates to a new method for assessing whether a		
CC	patient is afflicted with ovarian cancer. The method involves comparing		
CC	the expression level of a marker in a patient sample and the normal level		
CC	of expression of the marker in a control non-ovarian cancer sample, where		
CC	the marker is selected from 363 cancer markers described in the		
CC	specification. The method of the invention is useful in diagnosing or		
CC	characterising cancer, in detecting the presence of cancer as early as		
CC	possible, and the recurrence of ovarian cancer. The method may also be of		
CC	particular use with patients having an enhanced risk of developing		
CC	ovarian cancer (e.g. patients having a familial history of ovarian		
CC	cancer). The cancer markers may be used in the management and treatment		
CC	of e.g. brain and central nervous system disorders (e.g. bacterial and		
CC	viral meningitis, Alzheimer's disease or Parkinson's disease), brain		
CC	disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),		
CC	inflammations (e.g. bacterial or viral meningitis or encephalitis),		

	CC	testicular disorders (e.g. nonbuberculous granulomatous orchitis),		
	CC	connective tissue disorders, or heart disorders (e.g. ischemic heart		
	CC	disease or atherosclerosis). The compositions and methods may also be		
	CC	used in assessing the histological type of neoplasm associated with		
	CC	ovarian cancer, monitoring the progression of ovarian cancer, determining		
	CC	whether ovarian cancer has metastasized or is likely to metastasize,		
	CC	selecting a composition for inhibiting ovarian cancer, assessing the		
	CC	ovarian carcinogenic potential of a compound, or inhibiting ovarian		
	CC	cancer or at risk of developing ovarian cancer. The present amino acid		
	CC	sequence represents one of the ovarian cancer markers described in the		
	CC	invention		
XX				
SQ	Sequence 556 AA:			
Alignment Scores:				
Pred. No. :	1.03e-254	Length: 556		
Score:	2838.00	Matches: 556		
Percent Similarity:	100.00%	Conservative: 0		
Best Local Similarity:	100.00%	Mismatches: 0		
Query Match:	48.22%	Indels: 0		
DB:	5	Gaps: 0		
US-09-270-437D-8 (1-3283) x ABG96346 (1-556)				
OY	70	ATGATGAACAGCCTTTTAACTCGGGAAACCTGAGCCCCCGGTCAACGCCGAGCATTCGGG 125		
Dd	1	MetMetAsnLysLeuTyrILlegIyAsnLeuSerProAlaValThrLaAspAspLeuArg 20		
OY	130	CAGCTCTTTGGGGACAGGAAGTGCCCTCGGGGGACAGAGTCTCTGTAAGTCCGGCTAC 188		
Dd	21	GlnLeuPheGlyAspArgIysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40		
OY	190	GCCTTCGGGACTACCCCGACCAAGACTGGGCATCCGGCCATCGAGACCCTCGGGCT 249		
Dd	41	AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGlnThrLeuSerGly 60		
OY	250	AAAGTGAATTGCAATGGGAAAATCATGGAAGTTGATTCTCAAGTCTCTAAAAAGTAAAG 309		
Dd	61	LysValGlnLeuHisGlyLysIleMetGluValAspTyrSerValSerLysIysLeuArg 80		
OY	310	AGCAGAAAATTCAGATTGGAAAACATCCCTCTCACCTGCAGTGGAGGGTGTGGATGA 369		
Dd	81	SerArgLysIleGlnIleArgAsnIleProPohIsLeuGlnTrpGluValLeuAspGly 1000		
OY	370	CTTTTGGCTAATATGGGACAGTGGAGAAATGGGAAACAAGTCAACACACAGAAACC 429		
Dd	101	LeuLeuAlaGlnTyrGlyThrValGlnAsnValGlnGlnValAsnThrAspThrGlnTr 120		
OY	430	GCCGTTGTCAACGTCAATATGCCAACAGAAGAAAGCAAATAATGACATGGAGAAGCTA 489		
Dd	121	AlaValValAsnValThrTyrAlaThrArgGlnGlnAlaLysIleAlaMetGlnLysLeu 140		
OY	490	AGCGGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCCTACATCCCGGATGAAGAGTG 549		
Dd	141	SerGlyHisGlnPheGlnAsnTyrSerPheLysIleSerTyrIleProAspGlnGlnVal 160		
OY	550	AGCTCCCTTGGCCCCCTCAGGAGAGCCAGCGTGGGACCACTTTCCCAGAGCAAGGC 609		
Dd	161	SerSerProSerProProGlnArgAlaGlnAlaGlyAspHisSerSerArgGlnGlnGly 180		
OY	610	CACGCCCTCGGGGGCACTTCTCAGGCCACAGAAGATTGATTTCGCCGTGGAGATCCGCTGC 669		
Dd	181	HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuLysArgIleLeuVal 200		
OY	670	CCCAACCCAGTTTGTGTGCATCATCGGAAAAGAGGGCTTGACCATTAAGAACAACACT 729		
Dd	201	ProThrGlnPheValGlyAlaIleIleIleGlyLysGlnGlyLeuThrIleLysAsnIleThr 220		
OY	730	AAGCAGACCCAGTCCCGGAGTAGATTCATAGAAAAGAAGAACTTCGAGCTCAGAGAAG 789		
Dd	221	LysGlnThrArgInsSerArgValAspIleHisArgLysGlnAsnSerGlyAlaAlaGlnLys 240		
OY	790	CCTGTCAACATTCATGCGACCCCAAGAGGGGACCTTCTGAAGCATGCGGATGAATTCTGAA 849		


```

Db      241 ProValInHrIleHsAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY      850 ATCATGCAAGAAAGAGCGATGAGACCAACTAGCCGAGAGATTCTCTGAAAATCTTG 909
Db      261 IleMetGlnYsGluAlaAspGluThrIysLeuAlaGluGluIleProLeuYsIleLeu 280
QY      910 GCACAAATGGCTTGGTGGAAAGCTGATGGAAAAGAGGACGAAATTTAGAGAAATT 969
Db      281 AlaHsAsnGlyLeuValGlyArgLeuIleGlyYsGluGlyArgAsnLeuYsIle 300
QY      970 GAACATGAAACAGGGACCCCAATACATTCATCTTTGACGATTTGACATATACAC 1029
Db      301 GluHsGluThrGlyThrIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY      1030 CCGGAAAGAACCATCATCTGTGAGGGCAAGTTGAGGCTGTGCCAGTGTGAGATAG 1089
Db      321 ProGluArgThrIleThrValIleGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY      1090 ATTATGAAAGAGCTGCGTGAAGGCTTTGAAAATGATATGCTGGCTGTTAACACCACTCC 1149
Db      341 IleMetYsYsLeuArgGluAlaPheGluAsnMetLeuAlaValAsnThrIleSer 360
QY      1150 GGATATTCTCCAGGCTGTACCCCATCATCACAGTTTGCCGCTTCCGCATCATCACTCT 1209
Db      361 GlyTyrPheSerSerLeuTyrProHisHisGlnPheGlyProPheProHisHisSer 380
QY      1210 TATCCAGAGCGAGAGATTGTGATCTCTTCATCCCAACCCAGGCTGTGGGCCCATCATC 1269
Db      381 TyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAlaIleIle 400
QY      1270 GGGAGAAAGGGGGGACACATCAACAGCGGCGAGATTGCGCGAGCCTCTCAAGATT 1329
Db      401 GlyYsYsGlyAlaHisIleYsGlnLeuAlaAspPheAlaGlyAlaSerIleYsIle 420
QY      1330 GCCCTCGCGAGAGGCCCAAGCTGACCGAAGGATGGTCATCATCACCGGCGCACCGGAA 1389
Db      421 AlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyProProGlu 440
QY      1390 GCCCAGTTCAAGGCCCGAGGAGCGATCTTTGGGAAACTGAAAGAGGAAATCTTTTAC 1449
Db      441 AlaGlnPheYsAlaGlnIleArgIlePheGlyYsLeuYsGluGluAsnMetPheAsn 460
QY      1450 CCCAAGAGAGAGTAGAGCTGAGAGCGCATATCAGAGTCCCTCTCCACAGCTGGCCG 1509
Db      461 ProYsGluGlnValYsLeuGluAlaHisIleArgValProSerSerThrAlaGlyArg 480
QY      1510 GTGATTGGCAAGGTGGCAAGCCGTGAACGAACTGACAGACTTAAACAGTGCAGAAATC 1569
Db      481 ValIleGlyYsGlyGlyYsThrValAsnGluLeuGlnAsnLeuThrSerAlaGluVal 500
QY      1570 ATCGTGCCTCGTAGCAAAACGCCAGATGAAATAGAGAGTATCGTCAGATTTACGG 1629
Db      501 IleValProArgAspGlnThrProAspGluAsnGluValIleValArgIleIleGly 520
QY      1630 CACTTCTTTGCTAGACGACGCTGACAGCGCAGAGATCAGGAAATTTAACAAGGTGAAG 1689
Db      521 HisPhePheAlaSerGlnThrAlaGlnAlaGlyYsIleArgGluIleValGlnGlnValYs 540
QY      1690 CAGCAGAGACAGAAATACCTCTCAGGAGTGCCTCTCAGCGCAGCAAG 1737
Db      541 GlnGlnGlnGlnIleYsTyrProGlnGlyValAlaSerGlnArgSerYs 556

```

```

KW      Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX      food supplement; medical imaging; diagnostic; genetic disorder.
OS      Homo sapiens.
PN      WO200175067-A2.
PD      11-OCT-2001.
PE      30-MAR-2001; 2001WO-US008631.
PR      31-MAR-2000; 2000US-00540217.
XX      23-AUG-2000; 2000US-00649167.
PA      (HYSE-) HYSEQ INC.
PI      Dmanac RT, Liu C, Tang YT;
XX      WPI; 2001-639362/73.
DR      N-PSDB; AAS70981.
PT      New isolated polynucleotide and encoded polypeptides, useful in
PT      diagnostics, forensics, gene mapping, identification of mutations
PT      responsible for genetic disorders or other traits and to assess
PT      biodiversity.
PS      Claim 20; SEQ ID NO 37153; 103bp; English.
XX      The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC      sequences. (I) is useful as hybridisation probes, polymerase chain
CC      reaction (PCR) primers, oligomers, and for chromosome and gene mapping;
CC      and in recombinant production of (II). The polynucleotides are also used
CC      in diagnostics as expressed sequence tags for identifying expressed
CC      genes. (I) is useful in gene therapy techniques to restore normal
CC      activity of (II) or to treat disease states involving (II). (II) is
CC      useful for generating antibodies against it; detecting or quantitating a
CC      polypeptide in tissue, as molecular weight markers and as a food
CC      supplement. (II) and its binding partners are useful in medical imaging
CC      of sites expressing (II). (I) and (II) are useful for treating disorders
CC      involving aberrant protein expression or biological activity. The
CC      polypeptide and polynucleotide sequences have applications in
CC      diagnostics, forensics, gene mapping, identification of mutations
CC      responsible for genetic disorders or other traits to assess biodiversity
CC      and to produce other types of data and products dependent on DNA and
CC      amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC      amino acid sequences of the invention. Note: The sequence data for this
CC      patent did not appear in the printed specification, but was obtained in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences
XX      SQ      Sequence 614 AA:
XX
XX      Alignment Scores:
XX      Pred. No.: 1,5e-254 Length: 614
XX      Score: 2836.50 Matches: 567
XX      Percent Similarity: 92.05% Conservative: 0
XX      Best Local Similarity: 92.05% Mismatches: 4
XX      Query Match: 48.19% Indels: 45
XX      DB: 4 Gaps: 1

```

```

US-09-270-437D-8 (1-3283) x ABG06794 (1-614)
QY      23 GCCGGGTACCGGGCCCGGGAGACCCGCGCTCTCGGGGAGAGACGATGATGAACAAC 82
Db      1 AlaGlyTyrArgArg-ArgGluProArgPro-GlnGlyYsArgArgMetMetAsnYsIle 20
QY      83 TTATCATCGGGAACCTGAGACCCCGCTGACCGCGGACGACCTCGGGGAGCTTTGGGG 142
Db      20 euryIleGlyAsnLeuSerProAlaValThrAlaAspAspLeuArgGlnLeuPheCys 40
QY      143 ACAGGAAGCTGCCCTGCGGGGACAGATCTCTGAACTCGGCTACGCTTGTGAGCT 202
Db      40 spArgYsLeuProLeuAlaGlyGlnValLeuLeuYsSerGlyTyrAlaPheValAsp 60

```

```
QY 203 ACCCGACCAAGACTGGGCGCATCCGCCATCGAGACCCCTCGGGTAAAGTGAATTGC 262
Db 60 YRPRoAspGlnAsnThrPalaIleArgAlaIleGluThrLeuSerGlyysValGluLeuH 80
QY 263 ATGGGAAATCATGGAAGTTGATTACTGACTCTTAAAAAGCTAAAGAGCAGGAAATTC 322
Db 80 ISGlyLysIleMetGluValAspTyrSerValSerLysLeuArgSerArgLysIleG 100
QY 323 AGATTCCGAAACATCCCTCTCACTCCGAGTGGGAGSTTTGGATGGACTTTTGGCTCAAT 382
Db 100 InIleArgAsnIleProProHsiLeuGlnTyrGluValLeuAspGlyLeuLeuAlaGlnT 120
QY 383 ATGGGACAGTGGAGAAATGTGGAACAAGTCAACACAGACAGAGAAACCCCGTGTCAAG 442
Db 120 YRGIYThrValGluAsnValGluGlnValAsnThrAspThrGluThrAlaValAlaAsnV 140
QY 443 TCACATATGCAACAGAGAGAGAGCAAAATATGACCATGAGAGAGTAAAGCGGACATCAGT 502
Db 140 alHrrYrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSerGlyHsiGlnP 160
QY 503 TTGGAATCTACTCCCTTCAAGATTTCCTACATCCCGGAGAGAGAGTGAGCTCCCTTGC 562
Db 160 heGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluValSerSerProSerP 180
QY 563 CCCCTCAAGCAGCCAGCGTGGGAGCACTCTTCCCGGAGCAAGCGCCCGCTGGGG 622
Db 180 roProGlnArgAlaGlnArgGlyAspHsiSerSerArgGluGlnGlyHsiAlaProGlyG 200
QY 623 GCACCTTTCAGGCGCAACAGATTGATTTCCCGTGGAGATCTGGTCCCAACCCAGTTTG 682
Db 200 LyrThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValProThrGlnPheV 220
QY 683 TTGGTGCATCATCGGAGAGAGGGCTTGACCATTAAGAACATCATTAAGACGCCACT 742
Db 220 alGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLysGlnThrGlnS 240
QY 743 CCCGGGTAGATATCATAGAAAAGAAAGAACTGTGAGCTGCAGAGAGCGCTGCACATCC 802
Db 240 eHrrgValaPheIleHsiArgGlySerLysAsnSerGlyAlaAlaGluLysProValThrIleH 260
QY 803 ATGCCACCCCAAGAGGGGACTTTTGAAAGCATGCCGATGATTTCTTGAATCATGCAAGAG 862
Db 260 ISAlaThrProGluGlyThrSerGluAlaCysArgMetCileLeuGlnIleMetGlnLysG 280
QY 863 AGGAGATGAGACCAACTGATCCGGAAGAGATTCCTCTGAAAAATTTGGCACACATGGCT 922
Db 280 LuAlaAspGluThrLysLeuAlaGluGlnIleProLeuLysIleLeuAlaHsiAsnGlyL 300
QY 923 TGGTTGGAGACTGATTGAAAAGAGAGCAAAATTTGAAGAAAATTTGAACATGAACAG 982
Db 300 euValGlyArgLeuIleGlyLysGluGlyLysArgAsnLeuLysLysIleGlnHsiGluThrG 320
QY 983 GGAACCAAGATTAACATTTCAATCTTTTGAGAGATTTGACCATATACAAACCCGAGAAAGCA 1042
Db 320 LyrThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnProGluArgHrrI 340
QY 1043 TCACCTGGAGAGGCGACAGTGAAGCCCTGTGCCACTGTCTGGATGAGATTAAGAAGAC 1102
Db 340 IeThrValLysGlyThrValGluAlaCysAlaSerAlaGlnIleGlnIleMetLysIyl 360
QY 1103 TGGGTAGAGGCTTTGAAAATGATATGCTGGCTGTAAAC----- 1140
Db 360 euArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnGlnAlaAsnLeuIleProG 380
QY 1140 ----- 1140
Db 380 lYLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeuSerProProA 400
QY 1141 -----ACCACTCCGAGAT 1153
Db 400 laGlyProArgGlyAlaProProAlaAlaProTyrHsiProPheThrThrHsiSerGlyT 420
```

```
QY 1154 ACTTCTCAGGCTGTAAACCCCATCAGCAGTTTGGCCCGTTCCCGCATCATCTTTATC 1213
Db 420 YRPheserSerLeuTyrProHsiHsiGlnPheGlyProPheProHsiHsiSerTyrP 440
QY 1214 CAGAGCAGAGATTGTGAATCTTTCATCTCCACCCAGGCTGTGGGCGCCCATCATCGGGA 1273
Db 440 roGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAlaIleIleGlyL 460
QY 1274 AGAAGGGGGCACACATCAACAGCTGGCGAGATTGGCCGAGCCTTATCAAGATTGGCC 1333
Db 460 YslYsglyAlaHsiIleLysGlnLeuAlaArgPheAlaGlyAlaSerIleLysIleAlaP 480
QY 1334 CTGGCGAAGGCGCCAGAGCTCAGCGAAGAGATGTCATATACCGGGCGCCAGGAAGCC 1393
Db 480 roAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyProProGluAlaG 500
QY 1394 AGTTCAGGCGCCAGAGCGGATCTTTGGGAAACTGAAAGAGAGAAACTTCTTTAAACCCCA 1453
Db 500 InPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPheAsnProL 520
QY 1454 AAGAGAAAGTGAACTGTGAAGCGCATATCAGAGTGCCCTTTCACAGCTGGCCGGGTGA 1513
Db 520 YsgluGluValLysLeuGluAlaHsiIleArgValProSerSerThrAlaGlyArgValI 540
QY 1514 TTGGCAAAAGGTGGCAAGACCGGTGAACGACCTGCAAGACTTAACAGTGCAGAACTCATCG 1573
Db 540 IeGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAlaGlnValIleV 560
QY 1574 TGCTCTGTGACCAACCGCCAGATGAAATGAGAAAGTATCGTGAATTTATCGGGCACT 1633
Db 560 alProArgAspGlnThrProAspGluAsnGluGluValIleValArgIleIleGlyHsiP 580
QY 1634 TCTTTGCTAGCGACACTGTCACAGCGCAAGATCAGGGAATTTGACAAAGGTGAAGACG 1693
Db 580 hePheAlaSerGlnThrAlaGlnArgLysIleArgGlnIleValGlnGlnValLysGlnG 600
QY 1694 AGGAGCAAGAAATACCTCCAGGAGTCCGCTCAGACGCCAGCAAG 1737
Db 600 InGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 614

RESULT 4
AAU1613
ID AAU1613 standard; protein; 620 AA.
XX
AC AAU1613;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human novel secreted protein, Seq ID 1116.
KW Human; immunosuppressive; antiarthritic; antirheumatic; cyostatic;
KW cardiact; vasotrophic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnery;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
OS Homo sapiens.
XX
EN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
```

PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216889P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225577P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226272P.
PR 22-AUG-2000; 2000US-0226811P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236376P.
PR 29-SEP-2000; 2000US-0236388P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-488783/53.
XX DR N-PDB; AAS26150.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
PS Claim 11, SEQ ID NO 1116; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their

CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angioneu-
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Alignment Scores:

Pred. No.:	6,82e-252	Length:	620
Score:	2808.00	Matches:	560
Percent Similarity:	90.78%	Conservative:	1
Best Local Similarity:	90.61%	Mismatches:	11
Query Match:	47.71%	Indels:	46
DB:	4	Gaps:	2

US-09-270-437D-8 (1-3283) x ANU16163 (1-620)

```
QY 22 CGCGGCGTACCGGCGCGGCGGCGGCGCTTCGCGGGAAGACGG-----ATG 72
DB 3 ArgArgTyrAlaCysArgTyrArgSerGlyIleProGlySerThrHisAlaSerGlyMet 22
QY 73 ATGAACAGGCTTTACATCGGGAGCCTGAGCCCGCGCGTCACCGCGACAGCTCCGCGCAG 132
DB 23 MetGlnIysLeuArgIleGlyAsnLeuSerProAlaValThrAlaAspLeuArgGln 42
QY 133 CTCTTTGGGAGACAGAAAGCTGCCCCCTGGCGGAGCAGTCCGTGGAATCCGCGTACGCGC 192
DB 43 LeuPheGlyAspArgIleGlyLeuProLeuAlaGlyGlnValLeuLeuIysSerGlyTyrAla 62
QY 193 TTGCGTGAATACCCCGGACCGAAATGGGCGCATCCGCGCCATCGAGACCTTCGCGGTAA 252
DB 63 PheValAspTyrProAspGlnAsnThrAlaIleArgAlaIleGlnThrLeuSerGlyLys 82
QY 253 GTGGAATTGCATGGGAAATCATGGAGTGAATTACTCAGTCTTAAATAAGCTTAAGAGC 312
DB 83 ValGlnLeuHisGlyLysIleMetGlnValAspIysSerValSerIysLysLeuArgSer 102
QY 313 AGGAAATTCAGATTCCGAAACATCCCTCCTCACCCTGAGTGGAGGTGGTGAAGACTT 372
DB 103 ArgLysIleGlnIleArgAsnIleProProHisLeuGlnThrGlnValLeuAspGlyLeu 122
QY 373 TTGGCTCAATATGGGACAGTGGAGATGGGAAACAATCAACACAGACACAGAAACCGCC 432
DB 123 LeuAlaGlnTyrIleGlyThrValGluAsnValGlnGlnValAsnThrAspThrGlnThrAla 142
QY 433 GTTGTCAACGTCACATATGCAACAGAGAGAGAGAAATAATAGCATCGATGGAGACTTAAC 492
DB 143 ValValAsnValThrTyrAlaThrArgGlnGlnAlaLysIleAlaMetGlyLysLeuSer 162
QY 493 GGGCATCAGTTTGAAGACTACTCCTTCAGATTTCCTACATCCCGGATGAAGAGTGAAGC 552
DB 163 GlyHisGlnPheGlyAsnTyrSerPheLysIleSerTyrIleProAspGlnGlnValSer 182
QY 553 TCCGCTTCCGCGCGCTCAGCGAGCCCGAGCTGGGAGACCACTTTCGCGGAGCAAGGCGAC 612
DB 183 SerProSerProGlnArgAlaGlnArgIleAspHisSerSerArgGlnGlnGlyHis 202
```

```
QY 613 GCCCTGGGGGACATCTTTCAGGCGCAAGATGATTTCCGCTGCGGATCCTGTGCTCCC 672
DB 203 AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro 222
QY 673 ACCGAGTTGTTGGGCGCATCATCGGAAAGAGGGCTTGACCATTAAGAACTACTAG 732
DB 223 ThrGlnPheValGlyAlaIleIleIleGlySerGlnGlyLeuThrIleLysAsnIleThrLys 242
QY 733 CAGACCCAGTCCCGGCTGATATCATATCAATGAAAGAACTCGAGCTGCAGAGAGACCT 792
DB 243 GlnThrGlnSerArgValAlaAspIleHisArgLysGlnAsnSerGlyAlaAlaGlnLysPro 262
QY 793 GTACCATCATGCCACCCCGAGGAGACTTCTGAAGCATGCCGATGATCTTGAATATC 852
DB 263 ValThrIleHisAlaThrProGlnGlyThrSerGlnAlaCysArgMetIleLeuGlnIle 282
QY 853 ATGACGAAAGGCGAGAGAGAGAACCAATACCCGAAAGATTCCTCTGAAATCTTGCGCA 912
DB 283 MetGlnLysGlnAlaAspGlnThrLysLeuAlaGlnGlnIleProLeuLysIleLeuAla 302
QY 913 CACATAGGCTTGTTGGAAGACTGATTGGAAGAAAGAGCGAGAAATTGAAGAAATTGA 972
DB 303 HisAsnGlyLeuValGlyArgLeuIleGlyLysGlnGlyArgAsnLeuLysIleGln 322
QY 973 CATGAACAGGAGACCAAGATTAACAATCTCATCTTTGACAGATTTGAGCATATACACCG 1032
DB 323 HisGlnThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 342
QY 1033 GAAAGAACCATCATCTGGAAGAGGCGACGCTTGAGGCGCTGTCATGCTGCAATATGAGATT 1092
DB 343 GlnArgThrIleThrValLysGlyThrValGlnAlaCysAlaSerAlaGlnIleGlnIle 362
QY 1093 ATGAAGAGCTGCGTGAAGGCTTTGAAATGATATGCGCTGTATAC-----ACC 1140
DB 363 MetLysLysLeuArgGlnAlaPheGlnAsnAspMetLeuAlaValAsnGlnAlaAsn 382
QY 1140 ----- 1140
DB 383 LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 402
QY 1141 -----ACC 1143
DB 403 SerProProAlaGlyProAspArgIleAlaProProAlaAlaProTyrHisProPheThrThr 422
QY 1144 CACTCCGATATCTTCTCAGCTGTATCCCATCAACAGATTGGCCCGCTCCGCGATCAT 1203
DB 423 HisSerGlyTyrPheSerSerLeuTyrProHisHisGlnPheGlyProPheProHisHis 442
QY 1204 CACTCTTATCCAGACGAGAGATTGGAATCTCTTATCCCAACCCAGGCTGGGCGCC 1263
DB 443 HisSerTyrProGlnGlnGlnIleValAsnLeuPheIleProThrGlnAlaValGlyAla 462
QY 1264 ATCATGGGAAAGAGGGGCGACATCAACACTGCGCGAGATTCGCGGAGCTCTATC 1323
DB 463 IleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle 482
QY 1324 AAGATTGCCCTCGGAGAGGCCCGACAGCTCAAGGAAAGATGGTCATCATCAACCGGCGCA 1383
DB 483 LysIleAlaProAlaGlnGlyProAspValSerGlnArgMetValIleIleThrGlyPro 502
QY 1384 CCGGAAGCCCATGTTCAAGGCCCAAGGAGCGATCTTTGGGAAACTGAAGAGAAACTTC 1443
DB 503 ProGlnAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGlnGlnAsnPhe 522
QY 1444 TTTAACCCCAAAAGAGAGTGAAGCTGGAAGCGCATATCAAGATGCCCTTCCACAGCT 1503
DB 523 PheAsnProLysGlnGlnValLysLeuGln**HisIleArgValProSerSerThrAla 542
QY 1504 GCGCGGATGATTGGCAAGGTTGGCAAGACGTTGAAGAACTGAGAGAACTTAAACAGTGA 1563
DB 543 GlyArgValIleGlyLysGlyGlyLysThrValAsnGlnLeuGlnAsnLeuThrSerAla 582
```

QY 1564 GAAGTCATCGCTCGCTGACCAAGCAAGCAAGTGAAGTGTGCAAGT 1623
 DB 563 GlnValIleValProArgAspGlnThrProAspGlnGluValIleValArgIle 582
 QY 1624 ATCGGAGCACTCTTGTGACGACAGTGCACAGCGCAGATCAGGGAATTTGACCAAG 1663
 DB 583 IleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln 602
 QY 1684 GTGAAGCAGCAGGAGCAGAAATACCTTACGAGAGTCCGCTTACAGCCGACAGAG 1737
 DB 603 ValIysGlnGlnGlnGlnIleTyrProGlnGlyValAlaSerGlnArgSerLys 620
 RESULT 5
 ABUS5232
 ID ABUS5232 standard; protein, 620 AA.
 XX AC ABUS5232;
 XX DT 18-MAR-2003 (first entry)
 XX DE Human novel polypeptide #319.
 XX KW Human; neural disorder; immune system disorder; renal disorder;
 KW muscular disorder; respiratory disease; reproductive disorder;
 KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
 KW hyperproliferative disorder; inflammatory disease; allergic reaction;
 KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
 KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
 KW haemostatic; antiarteriosclerotic.
 XX OS Homo sapiens.
 XX EN US2002132753-A1.
 XX PD 19-SEP-2002.
 XX PF 17-JAN-2001; 2001US-00764864.
 XX PR 31-JAN-2000; 2000US-0179065P.
 XX PR 04-FEB-2000; 2000US-0180628P.
 XX PR 28-JUN-2000; 2000US-0214866P.
 XX PR 07-JUL-2000; 2000US-0216647P.
 XX PR 07-JUL-2000; 2000US-0216880P.
 XX PR 11-JUL-2000; 2000US-0217487P.
 XX PR 11-JUL-2000; 2000US-0217496P.
 XX PR 14-JUL-2000; 2000US-0218290P.
 XX PR 26-JUL-2000; 2000US-0220963P.
 XX PR 26-JUL-2000; 2000US-0224518P.
 XX PR 14-AUG-2000; 2000US-0224519P.
 XX PR 14-AUG-2000; 2000US-0225267P.
 XX PR 14-AUG-2000; 2000US-0225268P.
 XX PR 14-AUG-2000; 2000US-0225270P.
 XX PR 14-AUG-2000; 2000US-0225470P.
 XX PR 14-AUG-2000; 2000US-0225757P.
 XX PR 14-AUG-2000; 2000US-0225758P.
 XX PR 22-AUG-2000; 2000US-0226868P.
 XX PR 30-AUG-2000; 2000US-0228924P.
 XX PR 01-SEP-2000; 2000US-0229287P.
 XX PR 01-SEP-2000; 2000US-0229343P.
 XX PR 01-SEP-2000; 2000US-0229345P.
 XX PR 01-SEP-2000; 2000US-0229345P.
 XX PR 05-SEP-2000; 2000US-0229513P.
 XX PR 08-SEP-2000; 2000US-0231413P.
 XX PR 21-SEP-2000; 2000US-0234223P.
 XX PR 21-SEP-2000; 2000US-0234274P.
 XX PR 25-SEP-2000; 2000US-0234997P.
 XX PR 27-SEP-2000; 2000US-0235834P.
 XX PR 29-SEP-2000; 2000US-0236327P.
 XX PR 29-SEP-2000; 2000US-0236367P.
 XX PR 29-SEP-2000; 2000US-0236368P.
 XX PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239355P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249292P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX (ROSE/) ROSEN C A.
 PA (ROBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX Rosen CA, Ruben SM, Barash SC;
 XX WPI: 2003-147444/14.
 DR N-PSDB; ABX73491.
 XX New polypeptides and nucleic acids, useful in gene therapy for treating,
 PT inhibiting or preventing e.g. neural, immune system, muscular,
 PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
 PT renal disorders.
 XX Claim 11; SEQ ID NO 1116; 402pp; English.
 XX PS
 XX CC The invention relates to human novel polypeptides and their associated
 CC polynucleotides. The polypeptides and polynucleotides are useful in gene
 CC therapy for treating, inhibiting or preventing neural disorders, immune
 CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
 CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
 CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
 CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,
 CC (e.g. congenital heart defects, Epstein's anomaly and hypoplastic left
 CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
 CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
 CC leukemia), inflammatory diseases (e.g. septic shock, bursitis and
 CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
 CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABUS4914-ABUS5699 and
 CC ABUS5748 represent human novel polypeptides of the invention
 XX SQ Sequence 620 AA;
 Alignment Scores:
 Pred. No.: 6,82e-252 Length: 620
 Score: 2808.00 Matches: 560
 Percent Similarity: 90.78% Conservative: 11
 Best Local Similarity: 90.61% Mismatches: 11
 Query Match: 47.71% Indels: 46
 DB: 6 Gaps: 2
 US-09-270-437D-8 (1-3283) x ABUS5232 (1-620)
 QY 22 CGCCGGGTACCGGCGCGGAGCCGCGGCTCTCGGGAAGAGACGG-----ATG 72
 DB 3 ArgArgTyrAlaCysArgTyrArgSerGlyIleProGlySerThrHisAlaSerGlyMet 22
 QY 73 ATGAACAAGCTTATCATCGGGAACCTGAGCCCGCGCTCACCGCCGACGACCTTCGGCAG 132
 DB 23 MetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaSerPleuArgGln 42
 QY 133 CTCTTTGGGAGCAGGAAGCTGCTCCGCGGAGAGGTCTCTGCTGAAGTCCGGGCTAGGCC 192
 DB 43 LeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValIleLeuLysSerGlyTyrAla 62
 QY 193 TTCGTGACTACCCCGACGACGAACTGGGCCATCGCGCCCATCGAGACCCCTCGGGTAAA 252

Db 63 PheValAspTyrProAspGlnAsnTyrPalaIleArgGlaIleGluThrLeuSerGlyLys 82
QY 253 GTGGAATTCATGGGAAATCATGAGATTGATTACTGCTCTTAAAGCTAAGAGC 312
Db 83 ValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSer 102
QY 313 AGGAAATTCAGATTTCGAAACATCCCTCTCACTGACGTGGAGAGTGTTGGATGACT 372
Db 103 ArgLysIleGlnIleArgAsnIleProProHisLeuGlnIleThrGluValLeuAspGlyLeu 122
QY 373 TTGGCTCAATATGGGACATGGAGATGTGAAACAGTCAACACAGACAGAAACCCGC 432
Db 123 LeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAla 142
QY 433 GTTGCAACGTCACTATGCAACAGAGAAAGAAATAGCCATGAGAGAGCTTAAGC 492
Db 143 ValValAsnValThrTyrAlaThrArgGluGlnValLysIleAlaMetGluLysLeuSer 162
QY 493 GGGCATCAGTTTGAGAACTACTCTCTCAAGATTTCATCCGGATGAAAGAGGTGAGC 552
Db 163 GlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGlnValSer 182
QY 553 TCCCTCTGGCCCTCAAGCCGACGCGTGGGAGCACTCTCCGGAGCAAGGCCAC 612
Db 183 SerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGlyHis 202
QY 613 GCCCTGGGGGACATTCTCAGCGCAGACAGATTGATTTCCGCTGCCGATCTGTCCCC 672
Db 203 AlaProGlyGlyIleThrSerGlnAlaArgGlnIleAspPheProLeuArgGlyLeuValPro 222
QY 673 ACCGAGTTTGTGGTCATCATCGAAAGAGGGCTTGACCATTAAGAAATCATGACTAAG 732
Db 223 ThrGlnPheValGlyAlaIleIleGlyLysGlnLeuThrIleLysAsnIleThrLys 242
QY 733 CAGACCCAGTCCGGGTGATATCATGAAAGAGAACTCTGGAGCTGCAGAGAGCCT 792
Db 243 GlnThrGlnSerArgValAspIleHisArgLysGluLysSerGlyAlaIleGluLysPro 262
QY 793 GTACACCATGCATGCCACCCAGAGGGGACTTCTGAAGCATGCCAGATTTCTGAATC 852
Db 263 ValThrIleHisAlaThrProGlnGlyThrSerGlnLacysArgMetIleLeuGlnIle 282
QY 853 ATGAGAAAGAGGCGATGAGACCAACTAGCCGAGAGATTCTCTGAAATTTGGCA 912
Db 283 MetGlnLysGlnAlaAspGlnThrLysLeuAlaGlnIleIleProLeuLysIleLeuAla 302
QY 913 CACAATGGCTTGTGGAGAGCTGATTGAAAGAGAGAGCAAAATTTGAAGAAATTTGAA 972
Db 303 HisAsnGlyLeuValGlyArgLeuIleGlyLysGlnLysArgAsnLeuLysIleGlu 322
QY 973 CATGAACAGGGACCAAGATATCAATCTCTTTGGAGATTTGAGCATATACACCG 1032
Db 323 HisGlnThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 342
QY 1033 GAAAGAACCATCACTGAGAGGAGCAAGTTGAGGCGCTGCCAGTGTGATAGATGAT 1092
Db 343 GluArgThrIleLeuThrValLysGlyThrValGlnAlaCysAlaSerAlaGlnIleGluIle 362
QY 1093 ATGAAGAGCTGGGTGAGGCGCTTTGAAATGATATGCTGCTGTAAAC----- 1140
Db 363 MetLysLysLeuArgGlnAlaPheGlnAsnAspMetLeuAlaValAsnGlnGlnAlaAsn 382
QY 1140 ----- 1140
Db 383 LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 402
QY 1141 -----ACC 1143
Db 403 SerProProAlaGlyProArgGlyAlaProProAlaAlaAlaProTyrHisProPheThrThr 422
QY 1144 CATCTCGGATATCTTCTCAGCGCTGTACCCCATCACCAGTTTGGCCGCTTCCCGCATCAT 1203

Db 423 HisSerGlyTyrPheSerSerLeuTyrProHisIleGlnPheGlyProPheProHis 442
QY 1204 CACTTATTCAGACAGAGAGATTGTGAATCTTTCATCCAGCCAGGCTGTGGCGCC 1263
Db 443 HisSerTyrProGlnGlnGlnIleValAsnLeuPheIleProThrGlnAlaValGlyAla 462
QY 1264 ATCATCGGAGAAAGAGGGGCGACACTCAAAAGCTGGCGAGATTCGCGAGCTCTATC 1323
Db 463 IleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle 482
QY 1324 AAGATTGGCCCTGGCGAAAGGCCAGAGCTCAGCCGAAAGAGGTCATATCAACCGGCGCA 1383
Db 483 LysIleAlaProAlaGlnGlyProAspValSerGluArgMetValIleIleThrGlyPro 502
QY 1384 CCGAAGCCAGTTCAAGGCCAGGAGCGATCTTTGGGAAACTGAAGAGGAAACTTC 1443
Db 503 ProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGlnGlnAsn 522
QY 1444 TTTAACCCTCAAGAAAGAGTAAAGCTGAGAGCGCATATCAAGTGCCTTTCACAGCT 1503
Db 523 PheAsnProLysGlnGluValLysLeuGlu***HisIleArgValProSerSerThrAla 542
QY 1504 GCGCGGTGATTGGCAAGGTGAGCAAGCCGTGAACGAACTGCAAGACTTAACGAGTCA 1563
Db 543 GlyArgValIleGlyLysGlyLysThrValAsnGlnLeuGlnAsnLeuThrSerAla 562
QY 1564 GAAGTCATCGTCTGCTGTACCAACCGCCAGATGAAATGAGAAATGATCGTCAAGATT 1623
Db 563 GluValIleValProArgAspGlnThrProAspGlnLysGlnGlnValIleValArgIle 582
QY 1624 ATCGGAGCACTTCTTGTCTAGCCAGACTGCACAGCGCAAGATCAGGAAATTTGACACAG 1683
Db 583 IleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGlnIleValGlnGln 602
QY 1684 GTGAAGCAGCAGAGCAGAAATATACCTCAGGAGATCGCTCAGAGCCGAGAGAG 1737
Db 603 ValLysGlnGlnGlnGlnLysTyrProGlnGlyAlaIleAspGlnArgSerLys 620

RESULT 6
ID ABU89799 standard; protein: 555 AA.
XX
AC ABU89799;
XX
DT 10-JUL-2003 (first entry)
XX
DE Novel human protein NOV14a.
XX
KW Human; cytostatic; DAPK3-Agonist; DAPK3-Antagonist; cancer; NOV.
XX
OS Homo sapiens.
XX
PN WO2003031571-A2.
XX
PD 17-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031357.
XX
PR 05-OCT-2001; 2001US-0327454P.
PR 09-OCT-2001; 2001US-0327917P.
PR 09-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 25-NOV-2002; 2002US-0391342P.
PR 01-OCT-2002; 2002US-00262445.
XX
PA (CURA-) CTRAGEN CORP.


```
RESULT 7
AAV30649
ID AAV30649 standard; protein; 577 AA.
XX
AC AAV30649;
XX
DT 17-NOV-1999 (first entry)
XX
DE A murine c-myc coding region determinant binding protein.
XX
KW c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;
KW endonucleolytic attack; half-life; breast cancer; colon cancer;
KW pancreatic cancer.
XX
OS Mus musculus.
XX
PN M09946594-A2.
XX
PD 16-SEP-1999.
XX
PE 05-MAR-1999; 99WO-US004897.
XX
PR 09-MAR-1998; 98US-0077372P.
XX
PA (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI Ross J;
XX
DR WPI; 1999-551506/46.
DR N-PSDB; AAZ10617.
XX
PT Diagnosing presence or absence of a tumor in a human by examining c-myc
PT coding region determinant-binding protein.
PS
XX Example; Fig 1A-D; 79bp; English.
XX
CC The present sequence represents a murine c-myc coding region determinant
CC binding protein (CRD-BP). The presence or absence of a tumor can be
CC determined by determining the levels of CRD-BP present in the suspect
CC tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack
CC and so prolongs its half-life. The methods are used for diagnosing
CC presence or absence of a tumor in a human, especially breast, colon and
CC pancreatic cancer. They are also used to inhibit cancer cell growth
CC
XX
SQ Sequence 577 AA;
XX
Alignment Scores:
Pred. No.: 1.74e-170 Length: 577
Score: 1934.00 Matches: 381
Percent Similarity: 79.10% Conservative: 77
Best Local Similarity: 65.80% Mismatches: 95
Query Match: 32.86% Indels: 26
DB: 2 Gaps: 5
US-09-270-437D-8 (1-3283) x AAV30649 (1-577)
QY 73 ATGACAAAGGTTACATCGGGAAGTCCCGCGTCCACCGCAGACATTCGCGGAG 132
DB 1 MetAsnysLeuYrllleGlyAsnLeuAsnGlnSerValThrProAlaAspLeuGlnLys 20
QY 133 CTCCTTGGGAGACGAGAGCTGCCCTGGCGGACAGAGTCTGCTGAAGTCGGGCTACGCC 192
DB 21 ValPheAlaGlnHisIleSerIYrSerGlyGlnPheLeuValIYsSerGlyTYrAla 40
QY 193 TTCGTGACTACCCGACGACAACTGGGCGATCCGGCGCATCGAGACCTCTCGGGTAA 252
DB 41 PheValAspCysProAspGlnHisIleThrAlaMetIYsAlaIleGlnThrPheSerGlyLys 60
QY 253 GTGGAAATTCGATCGAGAAATCATGGAAGTTGATTAATCTCTTAAAGCTAAAGGAGAGC 312
DB 61 ValGlnLeuGlnIYsSarGlnLeuGlnMetGlnHisSerValProLysIYsGlnArySer 80
QY 313 AGGAAATTCAGATTGCAAAACATCCTCTCACCTGACATGGAGGTGTTGATGAGACTT 372
```

```
DB 81 ArgIYsIleGlnIleHAsnIleProGlnLeuAArgPGLValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAAGTCAACACAGACAAACCGCC 432
DB 101 LeuAlaGlnTYrGlyThrValGlnAsnCysGlnGlnValAsnThrGlnSerGlnThrAla 120
QY 433 GTTGTCAACGTCACATATGACAAAGAGAGACAAATAAGCCATGAGAGAGCTAAGC 492
DB 121 ValValAsnValThrTYrSerAsnArgGlnGlnThrArgGlnAlaIleMetIYsLeuAsn 140
QY 493 GGGCATCAGTTTGGAACTACTCCTTCAAGTTTCTTACATCCCGGATGAAGAGGTGAGC 552
DB 141 GlnIYsGlnLeuGlnAsnHisAlaLeuIYsValSerTYrIleProAspGlnGlnIleThr 160
QY 553 TCCGCTTCGCCCTCGCCGACGAGCCGAGCGTGGGACACATCTCCCGGAGCAAGCCAC 612
DB 161 -----GlnGlyProGlnAenGlyArgArgIYsGlyIYsPheGlySerArgGlyProArg 178
QY 613 -----GCCCTGGGGGCACTTTCAGGCCAGACAGATTGATTTCCCGCTG 657
DB 179 GlnGlySerProValAlaAlaGlyAlaProAlaIYsGlnGlnProValAspIleProLeu 198
QY 658 CGGATCCTGGTCCCGCACCGCATGTTGTGGCCATCATTCGGAAGAGGGGCTTACCATTA 717
DB 199 ArgLeuLeuValProThrGlnTYrValGlyAlaIleIleGlyLysGlnGlyAlaThrIle 218
QY 718 AAGAACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTGGA 777
DB 219 ArgAsnIleThrIYsGlnThrGlnSerIYsIleAspValHisSarGlyLysAsnAlaGly 238
QY 778 GCTGCAGAGAACCTGTACCATTCATCCACCCGACAGGGGACTTCTGAAGCATGGCCG 837
DB 239 AlaAlaGlnLysAlaIleSerValHisSerThrProGlnGlyCysSerSerIaCYLys 258
QY 838 ATGATTTGAAATTCATGAGAAAGAGCGAGATGAGACCAACTACGCCGAAGATTCCT 897
DB 259 MetIleLeuGlnIleMetHisIYsGlyAlaIYsAspThrIYsThrAlaAspGlnValPro 278
QY 898 CTGAAATCTTGGCACACATATGCTTGTGGAAGCTGATGGAAGAAAGAGCAGAAAT 957
DB 279 LeuIYsIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlnGlyArgAsn 298
QY 958 TTGAAGAAATTTGAACATGAACAGGACCAAGATTAACAATCTCATCTTTGACAGATTG 1017
DB 299 LeuIYsIYsValGlnGlnAspThrGlnThrIYsIleThrIleSerSerLeuGlnAspLeu 318
QY 1018 AGCATATPACAAACCGGAAAGAACCATCATCTGTAAGGCGACATTTAGGCGCTGGCCAGT 1077
DB 319 ThrLeuTYrAsnProGlnAArgThrIleThrValIYsGlyAlaIleGlnAsnCysCYsArg 338
QY 1078 GCTGAGATAGAGATTATGAAGAGCTGCGTGAAGCTTTGAAATGATATGATCGCTGTT 1137
DB 339 AlaGlnGlnGlnIleMetCysValArgGlnAlaIYsIYsGlnAsnAspValAlaAlaMet 358
QY 1138 AACACCATCTCC-----GATACCTTC----- 1158
DB 359 SerLeuGlnSerHisIleLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
QY 1159 ---TCAGCGCTGACCCCATCATCCAG-----TTTGGCCCGGTCCCGCAT 1200
DB 379 SerSerSerAlaValAlaProProProSerSerValThrGlyAlaAlaIaProTYrSerSer 398
QY 1201 CATCACTCTTATTCAGAGCAGAGATTTGATCTCTTCACTCCCAACCCAGGCTGGGCG 1260
DB 399 PheMetGlnAlaIaProGlnGlnIYsMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1261 GCCATATTCGGAGAAAGAGGGGCGACACATCAACACAGCTGGCGAGATTGCCGAGCTCT 1320
DB 419 AlaIleIleGlyIYsIYsGlyGlnHisIleIYsGlnLeuSerArgPheAlaSerAlaSer 438
QY 1321 ATCAAGATTGGCCCTCGCGGAAGCCGACGTCAGGAAAGATGTCTCATCTACCGGG 1380
```

Db 439 IleYstIleAlaPProPogInThrProAspSerLysValArgMetValIleThrGly 458
 QY 1381 CCACCGGAAGCCGATTCAGAGCCCGAGGAGCGATCTTTGGGAAACTGAAAGAGGAAAC 1440
 Db 459 ProPogInAlaInPheLysAlaGlnGlyArgIleTyGlyLysLeuYsGlnGluAsn 478
 QY 1441 TTCTTAAACCCCAAGAAAGAGTGAAGTGAAGCGCATTCAGAGTCCCTCTCCACA 1500
 Db 479 PhePheGlyProLysGlnGlnValLysLeuGlnThrHisIleArgValProAlaSerIle 498
 QY 1501 GCTGGCCGGTGAATGCGAAAGTGGCAAGACCGTGAACGAACTGCAAGCTTAACAGT 1560
 Db 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGlnLysLeuGlnAsnLeuThrAla 518
 QY 1561 GCAGAAAGTCACTGTCCTGTCGTCGCAAAACGCCAGTGAATGAAGAAAGTATGTCGCA 1620
 Db 519 AlaGlnValIleValProArgAspGlnThrProAspGlnAsnAspGlnValIleValLys 538
 QY 1621 ATTATCGGGCACTCTTTGCTAGCCAGACTGCAAGCCGCAAGATGAGGAAATTGTCA 1680
 Db 539 IleIleGlyHisPheTyAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558
 QY 1681 CAGGTGAAGCAGCAGCAGCAGCAAAATACCTCAGGAGTGCCTCAGCGCAGCAAG 1737
 Db 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgArgLys 577
 RESULT 8
 AAB11365
 ID AAB11365 standard; protein; 579 AA.
 AC AAB11365;
 XX 21-FEB-2001 (first entry)
 DE Human lung cancer associated antigen L523S.
 XX Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
 KM vaccine; detection.
 OS Homo sapiens.
 XX MO200061612-A2.
 PD 19-OCT-2000.
 XX 03-APR-2000; 2000MO-US008896.
 PF 02-APR-1999; 99US-00285479.
 PR 17-DEC-1999; 99US-00466396.
 PR 30-DEC-1999; 99US-00476496.
 PR 10-JAN-2000; 2000US-00480884.
 PR 22-FEB-2000; 2000US-00510376.
 XX (CORI-) CORIXA CORP.
 PA Wang T, Fan L;
 PI WPI; 2000-628399/60.
 DR N-PSDB; AAC66035.
 XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
 PT protein is used for detecting and monitoring progression of lung cancer
 PT in a patient.
 XX Claim 3; Page 259-261; 261pp; English.
 PS This invention describes a novel isolated polypeptide (I) which
 CC comprising an immunogenic portion of a lung tumor protein or variant (P2)
 CC which have cytostatic activity. The polypeptides and polynucleotides are
 CC used in compositions and vaccines to inhibit the development of cancer,
 CC especially lung cancer, in a patient. Methods described in the invention
 CC can be used to monitor the progression of a cancer by carrying out the
 CC detection at subsequent time points and comparing the results from the

CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
 CC are treated with P2, polynucleotides encoding P2 or antigen presenting
 CC cells expressing P2 and then administered to the patient to inhibit
 CC development of cancer

SQ Sequence 579 AA;

Alignment Scores:
 Pred. No.: 5,47e-165 Length: 579
 Score: 1875.00 Matches: 369
 Percent Similarity: 78.70% Conservative: 78
 Best Local Similarity: 64.96% Mismatches: 95
 Query Match: 31.86% Indels: 26
 DB: Gaps: 6

US-09-270-437D-8 (1-3283) x AAB11365 (1-579)

QY 73 ATGAACAAAGCTTTACATCGGGAAACCTGAGCCCCCGCCGACCCGAGACCTCCGAG 132
 Db 1 MetAsnLysLeuTyLleGlyAsnLeuSerGlnAsnAlaAlaProSerAspLeuGlnSer 20
 QY 133 CTCTTTGGGACAGGAAAGCTGCCCTGCGGGACAGTCTGCTGAATCCGGTACGCC 192
 Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyThrAla 40
 QY 193 TTCTGTGACTACCCCGACAGAACTGGGCACTCCGCGCATGTGAGACCTCTCGGTAA 252
 Db 41 PheValAspCysProAspGlnSerThrAlaLeuLysAlaIleGlnAlaLeuSerGlyLys 60
 QY 253 GTGCAATTCATGGGAAATATCATGGAATGTGATTCTCAGTCTCAAAAAGCTAGAGAGC 312
 Db 61 IleGlnLeuHisGlyLysProIleGlnValGlnHisSerValProLysArgGlnArgIle 80
 QY 313 AGGAATATTCAGATTCGAAACATCCCTCCTCAGTGCAGTGGAGAGTGTGATGACTT 372
 Db 81 ArgLysLeuGlnIleAlaArgAsnIleProHisLeuGlnIleProGlnValLeuAspSerLeu 100
 QY 373 TTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAATGTCAACAACAACAAGACCGCC 432
 Db 101 LeuValGlnTyGlyValIleGlnSerCysGlnGlnValAsnThrAspSerGlnThrAla 120
 QY 433 GTTGTCAACGTCACTATTCGAAACAAGAGAAAGAAAGAAATGATGAGAGTAAAGTAAAC 492
 Db 121 ValIleAsnValThrLysSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
 QY 493 GGGCATCAGTTTGAAGAACTACTCTTCAAGATTCTCCATCATCCCGATGAAGAAGTGAAC 552
 Db 141 GlyPheGlnLeuGlnAsnPheThrLeuLysValAlaTyIleProAspGlnThrAlaIle 160
 QY 553 TCCCTTCGCCCCCTCAG-----CGAGCCAGCGT-----GGGACCACTTTCCCGG 600
 Db 161 GlnGlnAsnProLeuGlnGlnProArgLysArgArgGlyLeuGlnLysGlnArgLysSer 180
 QY 601 GAGCAAGGCCACGCCCTGGGGCACTTCTCAGGCCAGCAAGATGATTTCCCGTGGCG 660
 Db 181 ArgGlnGly---SerProGlnSerValSerLysGlnLysProCysAspLeuProLeuAla 199
 QY 661 ATCCGTGTCGCCACCCAGTTTGTGTGTCATCATCGAAAGAGAGGAGCTTGACCATTAAG 720
 Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGlnGlnLysAlaThrIleArg 219
 QY 721 AACATCACTAAGCAGACCCGATCCGGGTAGATATTCATTAAGAAAGAACTGTGAGCT 780
 Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGlnAsnAlaGlyAla 239
 QY 781 GCAGAGAAGCTCTCCACCATGCATGCACCCAGAGGGAGCTTCAAGCATCCGCGATG 840
 Db 240 AlaGlnLysSerIleThrIleLeuSerThrProGlnLysThrSerAlaAlaCysLysSer 259
 QY 841 ATTCTTGAATCATGAGAAAGAGGAGCATGAGACCAACTAGCCGAAGAGATTCTCTG 900
 Db 260 IleLeuGlnIleMetHisLysGlnAlaGlnAspIleLysSerThrGlnGlnIleProLeu 279

```
QY 901 AAAATCTGGCACACATGGCTTGGTGGAGACTGATTGGAAAAAGAGCAGAAATTG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgSerLeu 299
QY 961 AAGAAATTTGAAACATGAAACAGGGACCAAGATTAACAATCTCATCTTTGGAGATTGGAC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY 1021 ATATGACACCGGAAAGAACATCACTGTGAAGGGCACAAGTTGAGCCTGTGGCTCT 1080
Db 320 LeuTyraAsnProGluIleThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATGAGATTATGACAGAGCTGCGTGAAGCTTTGAAAAATGATATGCTGCTTTAAC 1140
Db 340 GluGluGluIleMetLysLysIleArgLysSerTyrgluAsnAspIleAlaSerMetAsn 359
QY 1141 ACCCACTCC-----GGATACTTC----- 1158
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProProThr 379
QY 1159 -----TCCAGCCTGACCCCATCAACCGATTGGCCCTGCCGCAT 1200
Db 380 SerGlyMetProProProThrSerGlyProProSerAlaMetThrProProTyProGln 399
QY 1201 CATCACTCTTATCCAGACAGAGATTGTGAATCTCTTCATCCCAACCCAGCTGTGGCC 1260
Db 400 PheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeuSerValGly 418
QY 1261 GCCATCATCGGAAAGAGGGGGCACACATCAACAAGCTGGCGAGATTGGCGGACCTCT 1320
Db 419 AlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSer 438
QY 1321 ATCAAGATTGGCCCTGGGGAAGGCGCCAGCTGACCGAAAGATGTCATCATCAACCGGG 1380
Db 439 IleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIleIleThrGly 459
QY 1381 CCACCGGAGCCGAGTTCAAGGCCAGGACGAGATCTTTGGGAAATGAGAAAGAGAAAC 1440
Db 459 ProProGluIleGlnPheLysAlaGlnGlyArgIleTyrglyLysIleLysGluGlnAsn 478
QY 1441 TTCTTTAACCCCAAGAAAGATGAAGCTGAGACCGCATATCAGATGCCCTCTTCACA 1500
Db 479 PheValSerProLysGluGluValLysLeuGlnAlaHisIleArgValProSerPheAla 498
QY 1501 GCTGGCCGGGTGATTTGGCAAGGTGGCAGACCGTGAACCAACTCGGAATTACCACT 1560
Db 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuSerSer 518
QY 1561 GCAGAACTCATCGTGCCTCGTGAACCAACCGCAGATGAAATGAGAGATGATGTCAGA 1620
Db 519 AlaGluValValValProArgAspGlnThrProAspGluAsnAspGlnValValLys 538
QY 1621 ATTATGGGGCATCTTTGCTAGCCAGACTGCAACGCGCAAGATCAGGAAATTTGACAA 1680
Db 539 IleThrGlyHisPheTyraLysGlnValAlaGlnArgLysIleGlnGluIleLeuThr 558
QY 1681 CAGGTGAAGCAGAGAGAGAGAA 1704
Db 559 GlnValLysGlnHisGlnGlnGln 566
RESULT 9
ABB75053
ID ABB75053 standard; protein; 579 AA.
XX
AC ABB75053;
XX
DT 01-MAY-2002 (first entry)
XX
DE Human lung tumour LS2S3 recombinant protein sequence SPQ ID NO:446.
XX
KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX immune response.
XX
```

```
OS Homo sapiens.
XX
EN MO200200174-A2.
XX
PD 03-JAN-2002.
XX
PF 28-JUN-2001; 2001WO-US021065.
XX
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685626.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAM, Li SX, Kalos MD, Henderson RA,
PI McNeill PD, Fanger N, Retter MM, Marnerakis M, Fanger GR;
PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX
DR N-PSDB; ABL49297.
XX
FT WPI; 2002-090513/12.
XX
PS Polynucleotides encoding lung tumor polypeptides, useful for treating
PS lung cancer or stimulating an immune response.
XX
PS Claim 2; Page 365-367; 374pp; English.
XX
CC The present invention describes human lung tumour proteins. Human lung
CC tumour proteins and polynucleotides have cytostatic and immunostimulant
CC activities, and can be used in vaccine production. Compositions
CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC proteins, T cell populations, or antigen presenting cells that express
CC the lung tumour proteins are useful for treating lung cancer or
CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC ABB75070 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 579 AA;
XX
Alignment Scores:
Pred. No.: 5,47e-165 Length: 579
Score: 1875.00 Matches: 369
Percent Similarity: 78.70% Conservative: 78
Best Local Similarity: 64.96% Mismatches: 95
Query Match: 31.86% Indels: 26
DB: 5 Gaps: 6
US-09-270-437D-8 (1-3283) x ABB75053 (1-579)
QY 73 ATGAACAAGCTTACATCGGGAACCTGAGGCCCGCGGTACCGCGCACGACCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
QY 133 CTCCTTGGGACACGGAAGCTGCCCTGGCGGACAGATCCTGCTGAAGTCCGGCTACGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyraAla 40
QY 193 TTGCTGGACTACCCCGACCAAGAACTGGGCATCCGGCCCATGAGAACCTCTGGGGTAA 252
Db 41 PheValAspCysProAspGlnSerThrPalaLeuLysAlaIleGlnAlaLeuSerGlyLys 60
QY 253 GTGGAATTGCATGGGAAATCATGAAATGATTGATTCAGTCTCTAAAAGCTTAAGAGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AAGAAATTCAGATTGGAACAATCCCTCTACCTGACCTGAGTGGAGAGTGTGAGTGAATT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnIleProLysValLeuAspSerLeu 100
QY 373 TTGGCTCATATGAGGACACTGGAGAAATGTGGAACAAAGTCAACAGACACAGAAACCGCC 432
```


CC invention
 XX Sequence 579 AA;
 SQ Alignment Scores:
 Pred. No.: 5,47e-165 Length: 579
 Score: 1875.00 Matches: 369
 Percent Similarity: 78.70% Conservative: 78
 Best Local Similarity: 64.96% Mismatches: 95
 Query Match: 31.86% Indels: 26
 DB: 5 Gaps: 6
 US-09-270-437D-8 (1-3283) x ABB74997 (1-579)
 QY 73 ATGAACAAGCTTTTCATCGAGGACCTGAGCCCGCCGTCACCCGACGACCTCCGAGC 132
 Db 1 Metasnlyslleuylrileglyasnleuserglunsnalalaproseraspleuuser 20
 QY 133 CTCTTTGGGAGACAGAAAGCTGCCCCGCGAGACAGTCTGTGAAGTCCGAGTACGCC 192
 Db 21 llepelysaspalalysileprovalserglpropheleuvallystnrlgltyrAla 40
 QY 193 TTGCTGACATACCCGACACAGAACTGGGCCATCCGCGCCATGAGACCTCTCGAGTAA 252
 Db 41 PheValaspCysProaspGlnserTrpAlaleuylsAlaleglunalaleuSerGlylys 60
 QY 253 GTGGAATTCAGATGGGAAATCATGAAAGTTGATTACTACAGTCTTAAAGCTAAGAGC 312
 Db 61 lleglunelhsegllylrsproillegluvalghlsserValProlysArgGlnArgile 80
 QY 313 AGGAAATTCAGATTGGAACATCCCTCTACCTGAGTGGAGAGTGTGGATGACTT 372
 Db 81 ArglylserglunlleargasnilleProProhlsleuInlrpGlunalleuaspserleu 100
 QY 373 TTGGCTCAATATGGAGACGTGGAGATGTGAAACATCAACAGACACAGAAACCGCC 432
 Db 101 leuvalglntlyrlylvalValgluserCysglunlValasnThrAspserglunThrAla 120
 QY 433 GTTGTCAAGCTCATATGCAACAGAGAAAGCAAAATAGCCATGAGAAAGCTAAGC 492
 Db 121 ValValasnValThrlyrSerSerlylssaspGlnAlarglnAlaleuaspLySleasn 140
 QY 493 GGGCATCATGTTTGAAGACTACTCTTCAGATTTCCTACATCCCGGATGAAGAGTGGAGC 552
 Db 141 GlyPheglunleuGlunsnphethrleuylsValAlatylrilleProaspGlnThrAlaAla 160
 QY 553 TCCCTTCGCCCCCTCAG-----CGAGCCACGCT-----GGGACCACTTCCCGG 600
 Db 161 GlnGlnasnProleuGlnGlnProArgGlyArgArgGlyleuGlyGlnArgGlySerSer 180
 QY 601 GAGCAAGGCCACGCCCTGGGGGACATCTCTCAGGCCACAGACAGATTGATTTCCCGCTGGC 660
 Db 181 ArgGlnGly---SerProGlySerValSerlylsglnlyserProCysaspLeuProleuArg 199
 QY 661 ATCTGGTCCCCCAACCATGTTTGTGGTCATCATCGAAAGAGAGGCTTGACCATTAAG 720
 Db 200 leuLeuValProthrGlnpheValglYAlalleileglylsglnGlyAlamrilleArg 219
 QY 721 AACATACCTAAGAGACCCAGTCCCGGATGATATCCATGAAGAAAGAACTCTGAGCT 780
 Db 220 AsnilleThrlylsglnThrGlnserlylslaspValHlsargylsglnsnAlaglYAla 239
 QY 781 GGAGAGAACCTGTGACCATTCATCCACCCGAGAGGAGACTTCTGAAGATCCCGCAG 840
 Db 240 AlaGlnlylserlileThrilleuSerThrProGlnGlyThrSerAlaCyslylser 259
 QY 841 ATTCTTGAATCATGAGAAAGAGGACAGATGAGACCAAACTAGCCGAGAGATTCTCTG 900
 Db 260 lleleuGlnlleMetlslslysglnAlaglAspillelylserPheThrGlnGlnleProleu 279
 QY 901 AAAATTTGGCACACATGCTGTTGTTGAAGACTGATTGGAAAGAAAGGACGAAATTTG 960
 Db 280 LysilleleuAlahlsasnAsnphValglYArgyleuilleglylsglnGlylArglenleu 299

QY 961 AAGAAATTTGAACATGAAACAGGACCAAGATAACATCTCATCTTTCAGAGATTGAGC 1020
 Db 300 LysylsilleglunlAspThrAspThrlylserlileThrlylserProleuGlnleuThr 319
 QY 1021 ATATACAAACCCGGAAGAAACCATCATCTGTGAAGGACACAGTTCAGGCTGTGCCAGTCT 1080
 Db 320 leuTrpAsnProGlnArgThrilleThrVallylsglyAsnValglunThrCysAlalyAla 339
 QY 1081 GAGATAGAGATTATGAAAGACCTGCGAGGACCTTGAAGAAATGATATGCTGCTGTAC 1140
 Db 340 GlnGlnGlnleuMetlyllylslleArgGlnserlylrglnAsnAspilleAspMetasn 359
 QY 1141 ACCCACTCC-----GGATTACTTC----- 1158
 Db 360 leuGlnAlahlsleuilleProGlyleuAsnleuSnAlaleuGlyleuPheProProthr 379
 QY 1159 -----TCAGCTGTACCCCATCATCCAGTTTGCCCGCTCCCGCAT 1200
 Db 380 SerGlyMetProProProthrSerGlyProProserAlamethrProProtyrProGln 399
 QY 1201 CATCACTCTTATCCAGACAGAGATTTGTAATCTTCAATCCGACCCAGCTGTGGCC 1260
 Db 400 Pheglunl---SerGlnunThrGlnunThrValHlsleuPheilleProAlaleuSerValGly 418
 QY 1261 GCCATCATGCGGAAAGAGGCGGACACATCAACAGCTGGAGATTCGCGAGCCTCT 1320
 Db 419 AlallelleglylsglnGlnHlsillelylsglnleuSerArgPheAlaglYAlaser 438
 QY 1321 ATCAGATTTGCCCTCGCGGAAGGCCACAGCTCATGCGAAAGATGTGATATATACCGGG 1380
 Db 439 llelylslleAlaproAlaglunAlaproAspAlalyValArgmetValillelleThrGly 458
 QY 1381 CCACCGGAACCCAGTTTCAAGGCCACGAGCGATCTTGTGGAAACAGAAAGAGAAAC 1440
 Db 459 ProProGlnAlaglInPheylsAlaglnGlyArgillelyrGlyllylsglnGlnsn 478
 QY 1441 TTCTTTAACCCCAAGAAAGTGAAGCTTGAAGCGCATATCAGAGTGCCTCTCCACA 1500
 Db 479 PheValserProlylsglnGlnValylsleuGlnAlahlsilleArgValProserPheAla 498
 QY 1501 GCTGGCCGGGTGATTTGGCAAAAGTGGCAAGACCTGTGAACGAATGCGAACTTAACAGT 1560
 Db 499 AlaGlyArgValilleglylsglyllylsthValasnGlnleuGlnAsnleuSerSer 518
 QY 1561 GCAGAGTATCGTGCCTCGTGACCAACGCCGATGAAGAAATGAGAGATGCTGACA 1620
 Db 519 AlaGlnValValValProArgAspGlnThrProAspGlnsnAspGlnValVallyl 538
 QY 1621 ATTATCGGCACTTTCTTCTAGCCAGACTGCAACGCGCAAGATCAGAAATGTACAA 1680
 Db 539 lleThrGlyHlsPheThrAlaCysGlnValAlaGlnAlargylsilleGlnGlnleuThr 558
 QY 1681 CAGGTGAACAGACAGACGAGAA 1704
 Db 559 GlnVallylsglnHlsGlnGln 566
 RESULT 11
 ABB75054
 ID ABB75054 standard; protein; 579 AA.
 XX ABB75054;
 AC
 XX
 DT 01-MAY-2002 (first entry)
 XX
 DE Human lung tumour L523S recombinant protein sequence SEQ ID NO:449.
 KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
 KW immune response.
 XX
 OS Homo sapiens.
 XX
 PN MO200200174-A2.


```
Db 479 PheValSerProIysIleuGluValIysLeuGluAlaHisIleArgValProSerPheAla 498
QY 1501 GCTGGCCGGGTATTGGCAAGAGTGGCAGACCGGTGAAGAACTGCAAGAACTTAACCACT 1560
Db 499 AlaGlyArgValIleGlyIysGlyIysThrValAsnGluIleuGlnAsnLeuSer 518
QY 1561 GCAGAAATGATCGTGGCTGCTGACCAAGCCAGATGAAATGAAGAAATGATCGTCAGA 1620
Db 519 AlaGluValValProArgAspGlnThrProAspGluAsnAspGlnValValIys 538
QY 1621 ATTATCCGGCACTTCTTGTCTAGCCAGACTGCACAGCCGCAAGATAGAGAAATGTACAA 1680
Db 539 IleThrGlyHisPheIysValAcysGlnValAlaGlnArgIysIleGlnGluIleuThr 558
QY 1681 CAGGTGAAGCAGCAGCAGCAAAA 1704
Db 559 GlnValIysGlnHisGlnGlnGln 566

RESULT 12
ABP61917
ID ABP61917 standard; protein; 579 AA.
XX
AC ABP61917;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human lung cancer associated protein sequence SEQ ID NO:348.
XX
KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO200247534-A2.
XX
PD 20-JUN-2002.
XX
PF 30-NOV-2001; 2001WO-US047576.
XX
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA,
PI Monelli PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
DR WPI; 2002-583465/62.
XX
DR N-PSDB; ABQ92440.
XX
PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT the polynucleotides, useful in pharmaceutical compositions such as
PT vaccines and as markers to indicate the presence of lung cancer.
XX
PS Example 2; Page 337-339; 381pp; English.
XX
CC The present invention describes isolated human lung carcinoma
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC activity, and can be used in gene therapy and in vaccines. Compositions
CC comprising (I) or (II) can be used for stimulating an immune response in
CC a patient and for treating lung cancer in a patient. Oligonucleotides of
CC (I) can be used for detecting lung cancer in a patient. Oligonucleotides of
CC obtaining a biological sample from the patient, contacting the biological
CC sample with the oligonucleotide, detecting in the sample, an amount of
CC polynucleotide that hybridizes to the oligonucleotide and comparing the
CC amount of polynucleotide that hybridizes to the oligonucleotide to a
CC predetermined cut-off value, and determining the presence of a cancer in
CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
CC vaccines. (I) is useful as a marker to indicate the presence or absence
CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
CC ABP61992 represent sequences used in the exemplification of the present
```

```
CC invention
XX
SQ Sequence 579 AA;
Alignment Scores:
Pred. No.: 5,47e-165 Length: 579
Score: 1875.00 Matches: 369
Percent Similarity: 78.70% Conservative: 78
Best Local Similarity: 64.96% Mismatches: 95
Query Match: 31.86% Indels: 26
Dels: 5 Gaps: 6

US-09-270-437D-8 (1-3283) x ABP61917 (1-579)
QY 73 ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCCGTCACCGCAGCACTCCGGGCA 132
Db 1 MetAsnIysLeuIyrIleGlyAsnLeuSerGluAsnAlaIaProSerAspLeuIleuSer 20
QY 133 CTCTTTGGGGACAGGAAGCTGCTGCGGGAGCAGGTCTGCTGAGATCCGGCTAGCGC 192
Db 21 IlePheIysAspAlaIysIleProValSerGlyProPheIleuValIysThrGlyTyrAla 40
QY 193 TTGCTGACTACCCCGACAGAACTGCGGCGCATCCGCGCATCGAGACCCCTCCGGGTAA 252
Db 41 PheValAspCysProAspGlnSerTrpAlaLeuIysAlaIleGluAlaLeuSerGlyIys 60
QY 253 GTGAATTCATGCGAAATATCATGAAATTGATTACTGCTCTCTAAATGCTAAGAGAGC 312
Db 61 IleGluIleuHisGlyIysProIleGluValGluHisSerValProIysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCACTGCACTGCGGAGGTGTGATGACTT 372
Db 81 ArgIysLeuGlnIleArgAsnIleProHisIleuGlnTrpGluValIleuAspSerIleu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGATGGGAAACATGCAACAGACACAGCAACCGCC 432
Db 101 LeuValGlnIyrGlyValIleGluSerCysGlnIleuValAsnTrpAspSerGluThrAla 120
QY 433 GTTGTCAGCTCATATATGCAACAGAAAGCAAAATATGATGAGAGCTAAGC 492
Db 121 ValValAsnValIleThrIyrSerSerIysAspGlnAlaArgGlnAlaLeuAspIysLeuAsn 140
QY 493 GGGCATATGTTTGAAGACTACTCCTTCAAGATTTCCATATCCCGATGGAAGGTGAGC 552
Db 141 GlyPheGlnIleuGlnAsnPheIleuValAlaIleIyrIleProAspGluThrAlaAla 160
QY 553 TCCCTTGGCCCTCCTAG-----CGAGCCCAAGCT-----GGGACCACTTCCCGG 600
Db 161 GlnGluAsnProIleuGlnIleProArgIlyArgArgGlyIleuGlyIleuArgIysSer 180
QY 601 GACCAAGCCCAAGCCCTCGGGGCACTTCTCAGGCGCAGACAGATTGATTCGCGTGGG 660
Db 181 ArgGlnGlyI--SerProGlySerValSerIysGlnIysProCysAspLeuProIleuArg 199
QY 661 ATCTTGTCCTCCCAAGCTTTGTTGGTCCATATCATGGAAGAGAGGCTTGACCATAAAG 720
Db 200 LeuIleuValProThrGlnPheValIleGlyAlaIleIleGlyIysGlnGlyIleuIleArg 219
QY 721 AACATCACTAAGAGAGACCCAGTCCGGGTAGATATATCATGAAAGAAAGAACTCGAGCT 780
Db 220 AsnIleThrIysGlnIleuIleuSerIysIleAspValHisIleArgIysGluAsnAlaGlyAla 239
QY 781 GCAGAGAACCTGTCACTCATTCATGACCCAGAGGAGACTTCTGAAGATCGCGCATG 840
Db 240 AlaGluIysSerIleThrIleuSerThrProGluGlyThrSerAlaIacyIysSer 259
QY 841 ATTCTTGAATATCATGCAAGAAAGGAGCATGAGCAACCAATCAGCCAGAGATTCCTG 900
Db 260 IleIleuGluIleuMetHisIysIysIleuIleuAspIleIysPheThrGluIleuProIleu 279
QY 901 AAAATCTTGGCAGCAATATGCTTGGTGGAAACATGATTGAAAGAAAGAGCAATTTG 960
Db 280 IysIleIleuAlaHisAsnAsnIleValIleGlyArgIleuIleGlyIysGlnGlyArgAsnIleu 299
```



```

QY 961 AAGAAATTTGAACATGAACAGGAGCCAAAGATTAACATTCATCTTTGACGATTTGAGC 1020
DB 300 LyeIysIIeGIuGInAePTrAspThrLysIIeThrLysEserProLeuGInGluLeuThr 319
QY 1021 ATATACAAACCCGAAACACCATGTCAGTGAAGGCAACAGTTGAGGCGCTGACGAGTCT 1080
DB 320 LeuYrAsnProGInAeGTrHrIleThrValLysGlnValGlnThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTTATGAAGAGAGCTGTCAGGAGCTTTGAAAATGATGATGCTGTTAAC 1140
DB 340 GluIuGInIleMetLysLysIIeArgGlnSerYrGlnAsnAePTrIleAlaSerMetAsn 359
QY 1141 ACCCACTCC-----GGATCTTC----- 1158
DB 360 LeuGlnAlaHIsLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProProThr 379
QY 1159 -----TCCAGCGCTTACCCCATCACAGTTTGCGCCGTCAT 1200
DB 380 SerGlyMetProProProThrSerGlyProProSerAlaMetThrProProTyProGln 399
QY 1201 CATCACTCTTATCCAGCAGAGATTTGTGATCTTTTCATCCCAACCCAGGCTGTGGCC 1260
DB 400 PheGluGIn-----SerGluThrGluThrValHIsLeuPheIleProAlaLeuSerValGly 418
QY 1261 GCCATCATCGGAGAGAGGCGGCAACATCAACAGCTGGCGAGATTGGCGGAGCCGCT 1320
DB 419 AlaIleIleGlyLysGlnGlyGlnHIsIleLysGlnLeuSerArgPheAlaGlyAlaSer 438
QY 1321 ATCAAGATTTGCGCCGCGGAAAGCCAGACGTCAGCGAAAGGATGTCATCATCCCGG 1380
DB 439 IleLysIleAlaProAlaGlnAlaProAspAlaLysValArgMetValIleIleThrGly 458
QY 1381 CCACCGGAGCCAGCTTCAAGGCGCCAGGAGCGGATCTTTGGAAAAGTGAAGAGGAAAC 1440
DB 459 ProProGlnAlaGlnPheLysAlaGlnGlyArgIleYrGlyLysIleLysGlnGlnAsn 478
QY 1441 TTCTTTAACCCTAAAGAAAGATGAAGCTGGAAGCGCAATCAAGAGTGCCTTCCACA 1500
DB 479 PheValSerProLysGlnIleValLysLeuGlnAlaHIsIleGlyAlaProSerPheAla 498
QY 1501 GGTGCGCGGATGTTGGCAAGGTGCGCAAGCGTGAACGAACTGCAAGATTTAACAGT 1560
DB 499 AlaGlyAlaValIleGlyLysGlyLysThrValAsnGlnLeuGlnAsnLeuSerSer 518
QY 1561 GCAGAAGTATCGTGCCTGCTGTCACCAACGCCAGATGAAGAAATGAGGAAGTATCTCA 1620
DB 519 AlaGlnValValAlaProArgAspGlnThrProAspGlnAsnAspGlnValValLys 538
QY 1621 ATTATCGGCACTTTCTTGTAGCCAGACTGCAACCGCAAGTTCAGGAAATGTACAA 1680
DB 539 IleThrGlyHIsPheYrAlaCysGlnValAlaGlnArgLysIleGlnGlnIleLeuThr 558
QY 1681 CAGGTGAAGCAGCAGGAGAGAA 1704
DB 559 GlnValLysGlnHIsGlnGlnGln 566

```

RESULT 13
ABP61974 standard; protein; 579 AA.

```

XX AC ABP61974;
XX DT 07-OCT-2002 (first entry)
XX DE Human lung cancer associated protein sequence SEQ ID NO:449.
XX KW Human; lung cancer; lung tumour; cyrostatic; gene therapy; vaccine.
XX OS Homo sapiens.
XX PN MO200247534-A2.
XX

```

```

PD 20-JUN-2002.
PF 30-NOV-2001; 2001WO-US047576.
XX 12-DEC-2000; 2000US-00735705.
XX 07-MAY-2001; 2001US-00850716.
XX 28-JUN-2001; 2001US-00897778.
XX
XX (CORI-) CORIXA CORP.
XX Wang T, Wang A, Szelky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
XX Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX WPI; 2002-583465/62.
XX N-PSDB; ABQ92485.
XX
XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
XX the polynucleotides, useful in pharmaceutical compositions such as
XX vaccines and as markers to indicate the presence of lung cancer.
XX
XX Claim 9; Page 375-377; 381pp; English.
XX
XX The present invention describes isolated human lung carcinoma
XX polynucleotides (I) and polypeptides (II). (I) and (II) have cyrostatic
XX activity, and can be used in gene therapy and in vaccines. Compositions
XX comprising (I) or (II) can be used for stimulating an immune response of
XX a patient and for treating lung cancer in a patient. Oligonucleotides of
XX (I) can be used for detecting the presence of a cancer in a patient, by
XX obtaining a biological sample from the patient, contacting the biological
XX sample with the oligonucleotide, detecting in the sample, an amount of
XX polynucleotide that hybridises to the oligonucleotide and comparing the
XX amount of polynucleotide that hybridises to the oligonucleotide to a
XX predetermined cut-off value, and determining the presence of a cancer in
XX the patient. (II) and (II) are useful in pharmaceutical compositions, e.g.
XX vaccines. (I) is useful as a marker to indicate the presence or absence
XX of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
XX ABP61992 represent sequences used in the exemplification of the present
XX invention
XX
XX Sequence 579 AA:
XX
XX Alignment Scores:
XX Pred. No.: 5,47e-165 Length: 579
XX Score: 1875.00 Matches: 369
XX Percent Similarity: 78.70% Conservative: 78
XX Best Local Similarity: 64.96% Mismatches: 95
XX Query Match: 31.86% Indels: 26
XX DB: Gaps: 6
XX
XX US-09-270-437d-8 (1-3283) x ABP61974 (1-579)
XX
XX 73 ATGAACAGCTTATACATCGGAGACCTGAGCCCGCGTACCGCGAGACCTCGGGCAG 132
XX 1 MetAsnLysLeuYrIleGlyAsnLeuSerGlnAsnAlaIleProSerAspLeuGlnSer 20
XX 133 CTCCTTGGGAGCAGAGAGCTGCCCTGGCGGAGAGCTCTCTGCTGAAGTCCGCTACGCC 192
XX 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
XX 193 TTCGTGACTACCCGACGACCAAGCTGCGGCATCGCGGCATCGAGACCTCGGGTAAA 252
XX 41 PheValAspCysProAspGlnSerIlePheValAlaIleGlnAlaLeuSerGlyLys 60
XX 253 GTGGAATTGCAATGGGAAATCATGAGATGTACTCACTCTTAAAGCTAAGAGAGC 312
XX 61 IleGlnLeuHIsGlyLysProIleGlnValGlnHIsSerValProLysArgGlnArgIle 80
XX 313 AGGAAATTCAGATTGCAACATCCCTCTCACTGAGTGGAGGAGTGTGATGACTT 372
XX 81 ArgLysLeuGlnIleArgAsnIleProProHIsLeuGlnThrGlnValLeuAspSerLeu 100
XX 373 TTGGCTCAATATGAGGACAGTGGAGAAATGTGGAACAAGTCAACACAGAACCGGCC 432

```

Db 101 LeuValGlnTyrGlyValValGlnSerCysGlnGlnValAsnThrAspSerGlnThrAla 120
QY 433 GTTGTCAAGTCACATATGCAACAAGAGAGAAAGAAAATGCCATGAGAGAACTAAC 492
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAAGAACTCTCTTCAAGATTTCTCATCTCCGGATGAAGAGTGAAC 552
Db 141 GlyPheGlnLeuGlnAsnPheThrLeuLysValAlaTyrIleProAspGlnThrAlaAla 160
QY 553 TCCCTTGGCCCTCAG-----CGAGCCCAAGCGT-----GGGAGCACTCTTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAAGGCGACGCCCCCTGGGGGCACTTCTCAAGGCGACAGATTGATTTCCCGCTGG 660
Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGTGATCCCAACCCAGTTTGTGTGTCATCATCGGAAAAGAGGGCTTGAACAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGlyGlyAlaThrIleArg 219
QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTGTGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGlnAsnAlaGlyAla 239
QY 781 GCAGAGAAGCTGTGCACCATCATGCCCCAGAGGGAGATTTCTGAAGCATGCCGACATG 840
Db 240 AlaGlnLysSerIleThrIleLeuSerThrProGlnGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAATCATGACGAAAGAGAGAGAGATGAGACCAACTAGCCGAGAGATTCCTGTG 900
Db 260 IleLeuGlnIleMetHisLysGlyAlaGlnAspIleLysPheThrGlnGlnIleProLeu 279
QY 901 AAAATCTTGGACACCAATGCGTTGGTGAAGACTGATGGAAAAGAGAGCAAAATTTG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlnGlyArgAsnLeu 299
QY 961 AAGAAATGGAACATGAACAGGAGCAACAGTAATCATCTTGTTCGAGAGATTTAGAC 1020
Db 300 LysLysIleGlnGlnAspThrAspThrLysIleThrIleSerProLeuGlnGlnLeuThr 319
QY 1021 ATATACAAACCCGGAAGAACCATCATCTGTGAAGGGACAGTGAAGGCTGTGACAGCT 1080
Db 320 LeuTyrAsnProGlnArgThrIleThrValLysGlyAsnValGlnThrCysAlaLysAla 339
QY 1081 GAGATGAGATTTAGAAAGACTGCGTGAAGCCCTTTGAAAATGATATATGCTGTGAAC 1140
Db 340 GlnGlnGlnIleIleMetLysLysIleArgGlnSerTyrGlnAsnAspIleAlaSerMetAsn 359
QY 1141 ACCCACTCC-----GGATTCCTTC-----GGAATCCTTC----- 1158
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProThr 379
QY 1159 -----TCCAGCTGTGATCCCGCATCACAGTTTGGCCGCTCCCGCAT 1200
Db 380 SerGlyMetProProThrProThrSerGlyProProSerAlaMetThrProPolyProGln 399
QY 1201 CATCACTCTTATCCAGAGCAGAGATTGTAATCTTCTTCAACCCAGAGCTGTGGC 1260
Db 400 PheGlnGln---SerGlnThrGlnThrValHisLeuPheIleProAlaLeuSerValGly 418
QY 1261 GCCATATGGGAGAGAGGGGGGACACATTAACAGCTGGCAGATTCGCCGAGCTCTT 1320
Db 419 AlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSer 438
QY 1321 ATCAAGATTGCCCTCGGAGAGGCCAGACGTCAAGGAAAGATGTGCATCAACCGGG 1380
Db 439 IleTyrIleAlaProAlaGlnAlaProAlaValAlaArgPheValIleIleThrGly 458
QY 1381 CCACCGGAAGCCCAAGTTCAAGGCCCAAGGAGCGGATTTTGGAAACTGAAAGAGAAAC 1440

Db 459 ProProGlnAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLysGlnGlnAsn 478
QY 1441 TTTCTTAACCCCAAGAAAGAGTGAAGCTGGAAGGCGATATGACAGTCCCTTCCACA 1500
Db 479 PheValSerProLysGlnGlnValLysLeuGlnAlaHisIleArgValProSerPheAla 498
QY 1501 GCTGGCCGGGTGATTTGGGCAAAAGGTGGCAAGCCGTGAACGAACTGCAACTTAACAGT 1560
Db 499 AlaGlyArgValIleIleGlyLysGlyGlyLysThrValAsnGlnLeuGlnAsnLeuSerSer 518
QY 1561 GCAGAAATCATCGTCCCTCGTGAACGCAACGCGACATGAAATGAGAAAGTGAATCGTACA 1620
Db 519 AlaGlnValValAlaProArgAspGlnThrProAspGlnAsnAspGlnValValLys 538
QY 1621 ATTATCGGACACTTCTTGTCTAGCCAGACTCAACAGCCCAAGATAGAGGAAATGTACA 1680
Db 539 IleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlnIleLeuThr 558
QY 1681 CAGGTGAAGCGACGACAGACGAAA 1704
Db 559 GlnValLysGlnHisGlnGlnGln 566

RESULT 14
ABP61973
ID ABP61973 standard; protein, 579 AA.
AC ABP61973;
XX 07-OCT-2002 (first entry)
DT
DT
XX Human lung cancer associated protein sequence SEQ ID NO:446.
DE Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
XX
XX Homo sapiens.
OS
XX W0200247534-A2.
FN
XX 20-JUN-2002.
PD
XX 30-NOV-2001; 2001WO-US047576.
PF
XX 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI McNeill PD, Fanger N, Rietter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX
XX WPI; 2002-583465/62.
DR N-PSDB; ABQ92483.
XX
XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
PT the polynucleotides, useful in pharmaceutical compositions such as
PT vaccines and as markers to indicate the presence of lung cancer.
XX
XX Claim 9; Page 372-374; 381pp; English.

CC The present invention describes isolated human lung carcinoma
CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
CC activity, and can be used in gene therapy and in vaccines. Compositions
CC comprising (I) or (II) can be used for stimulating an immune response in
CC a patient and for treating lung cancer in a patient. Oligonucleotides of
CC (I) can be used for detecting the presence of a cancer in a patient, by
CC obtaining a biological sample from the patient, contacting the biological
CC sample with the oligonucleotide, detecting in the sample, an amount of
CC polynucleotide that hybridises to the oligonucleotide and comparing the
CC amount of polynucleotide that hybridises to the oligonucleotide to a
CC predetermined cut-off value, and determining the presence of a cancer in
CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.

CC vaccines. (1) is useful as a marker to indicate the presence or absence
 CC of a cancer such as lung cancer. AB092145 to AB092486 and ABP61866 to
 CC ABP61992 represent sequences used in the exemplification of the present
 CC invention

XX Sequence 579 AA:

Alignment Scores:

Pred. No.:	5,47e-165	Length:	579
Score:	1875.00	Matches:	369
Percent Similarity:	78.70%	Conservative:	78
Best Local Similarity:	64.96%	Mismatches:	95
Query Match:	31.86%	Indels:	26
DB:	5	Gaps:	6

US-09-270-437d-8 (1-3283) x ABP61973 (1-579)

```

QY 73 ATGAACAGCTTTATACATCGGGAACTGAGCCCGCCGTCACCGCCGAGACCTCCGCGCAG 132
DB 1 Metasnllyleutrylllelyasnleusergluamalaalaproseraspleuoluser 20
QY 133 CTCTTTGGGAGCAGAGAGCTGCGCTGGCGGAGCAGTCTCTGAAAGTCCGGCTACGCG 192
DB 21 llephelyaspaalalyseillevrovalyserglipropheleuvalylslythrylalyala 40
QY 193 TTCGTGAGTACCCCGCAGCAGAACTGGCGCATCCCGGCGCATGAGACCTCTCGGGTAA 252
DB 41 Phevalaspcysproaspluserltrialaleuylsalaleolualaleusergllyls 60
QY 253 GTGGAATTCATGGAAGAAATCATGGAAGTGTACTGAGTCTTAAAGAAAGTAAAGAGC 312
DB 61 lleglulehhslelylsprolleglvalgluhislerValProlysarlglnhrglyle 80
QY 313 AGGAAATTCAGATTGGAACATCCCTCTCATCTGCACTGGAGGCTTTGGATGAGACTT 372
DB 81 Arglylsleuglnllearghsnllleprohlsleuglnltpgluvalleuhspsereu 100
QY 373 TTGGCTCAATTATGGGACAGTGGAGATGTGGAAAGTCAACACAGACAGAAACCGCC 432
DB 101 leuvalglntlylvalvalglusercysgluvalasnthraspergluthrala 120
QY 433 GTTGTACAGCTCATATGCAACAGACAGAAAGAAATAGCCATGAGAAAGCTAACG 492
DB 121 Valvalasnlvalthrylserlserlysaspglnalalarglnlaleuhsprylsleuan 140
QY 493 GGGCATCAGTTGAGAACTACTCTTCAAGATTCTCTACATCCCGGATGAAAGGTGAGC 552
DB 141 GlyPheGlnleuglnuhsnphethrleuylsValalalyrleproaspgluthralaala 160
QY 553 TCCCTTCGCGCCCTCAG-----CGAGCCAGCGT-----GGGAGCAGCTCTTCCCGG 600
DB 161 GlnGlnasnProleuglnnglnProargglyarghrglyleuylglnhrglylserSer 180
QY 601 GAGCAAGGCGACCGCTCGGGGCACTTTCAGGCGCAGACAGATGATTTCCGCTGCGG 660
DB 181 Arglnslyl---serprogllyserValserlysglnlylserProcyaspleuProleuArg 199
QY 661 ATCTGTGTCACCAACCGATTGTGTGTGTCATCATCGGAAAGAGGCTTGACATTAAG 720
DB 200 leuvalvalProthrghnphelvalgllyalalellelylsglnlylathnlllearg 219
QY 721 AACATCCTAAGAGACCGCCGCTCGGGTAGATTCATGAGAAAGAAAGAACTGTGAGCT 780
DB 220 Asnlethrlysglnthnglnserlyslleaspaalhlsaglylsglnasnlalaglyala 239
QY 781 GCAGAGAGCGCTGCACATCATGACCAACCCAGAGGAGACTTCTGAAGCATGCGCATG 840
DB 240 AlaagllylserllyethrilleuserThrProgluglylthrserralalacyllylser 259
QY 841 ATTCTTGAATCATGCAAGAGAGCGCAGATGAGACAAACTAGCCGAAAGATTCCTCTG 900
DB 260 lleuuglnlleuethlslysglnalaglnaspllelyhsphenhrglnhrglnleProleu 279

```

```

QY 901 AAATCTTGACACAATAGCTGTGGAGAGCTGATGGAAAAAGAGCGAAATTTG 960
DB 280 lyslleuvalahlsasnasnphelvalgllyarghrglylserlylsglnlyllyarghsnleu 299
QY 961 AAGAAATTTGAACTGAAACAGGAGCAAGATTAACATCTCATCTTTGACAGATTGAGC 1020
DB 300 lysylslelgluhsnphthrasphrlysllethrlylserProleuhsnleuhsn 319
QY 1021 ATATACAAACCCGAGAAAGAACCATCATCTGTGAAGGCGACAGTGTAGGCTGTGCCAGTCT 1080
DB 320 leuhyrhanpProgluargThrillethrVallylsglnvalgluthrCysalalyala 339
QY 1081 GAGATGAGATTTATGAGAGAGCTGCGTGAAGCCTTTGAAATGATGATGCTGGCTTAC 1140
DB 340 Gluuglnlileuethlylserlylserlylserlylserlylserlylserlylserlylser 359
QY 1141 ACCACCTCC-----GGATACCTC----- 1158
DB 360 leuGlnalAhlsleuThlProgllyleuhsnleuamalaaleuyllyleuhsnProthr 379
QY 1159 -----TCCAGCCTGTACCCCGCATACCAAGTTGGCCCGCTCCCGCAT 1200
DB 380 SerGlyMetProProProThrserylProProseralameThrProProtyrProgin 399
QY 1201 CATCACTCTTATCCAGACAGAGAGATGTGTGATCTCTTATCCCAACCAAGGCTGTGAGC 1260
DB 400 Pheglugln---SerGlnthrgluthrvalhlsleuhsnleuhsnleuhsnleuhsnleuhsn 418
QY 1261 GCCATCATCGGGAGAGAGGCGGACACATCAACAGCTGGCGAGATTCGCGGAGCTCT 1320
DB 419 Alallellelylylsglnlylshlslelylserlylserlylserlylserlylserlylser 438
QY 1321 ATCAAGATTCGCCCTCGGAGAGCGCCAGACGTGACGAAAGAGATGTCTATCAACCGGG 1380
DB 439 llelysllealProalaglnalProalalyValargMetVallellethrgly 458
QY 1381 CCACCGAAGCCCGATTTCAAGGCCCGACGAGATCTTTGGAAACTGAAAGAGAAAC 1440
DB 459 Proproglnalaglnhphelysalainglylserlylserlylserlylserlylserlylser 478
QY 1441 TTCTTTAACCCCAAGAAAGATGAAAGCTGAAAGCCATATCAGATGCCCTTCCACA 1500
DB 479 PhevalserProlysglnlvallylserlylserlylserlylserlylserlylserlylser 498
QY 1501 GCTGGCGGGTATTTGGCAAGGTGCAAGACCGTGAACGATCTGCAAGAACTTAAACAGT 1560
DB 499 AlaGlyArgValillelylserlylserlylserlylserlylserlylserlylserlylser 518
QY 1561 GCAGAACTCATGCTGCTGCTGACCAAGCGCAGATGAAAGATGAGAAAGATGATGATGATGAT 1620
DB 519 AlaGlnvalValalProarghrglnthnProaspglnuhsnspglvalvallyls 538
QY 1621 ATTATGCGGCACTTCTTGTAGCAGACTGCAAGCGCAAGATCAAGGAAATTTGACAA 1680
DB 539 llethrGlyhlsheyrhalaCysglnvalalaglnhrglylserlylserlylserlylserlylser 558
QY 1681 CAGGTGAAGCAGAGACAGAAA 1704
DB 559 Glnvallylserlylserlylserlylserlylserlylserlylserlylserlylserlylser 566

```

RESULT 15
 ADA28536
 ID ADA28536 standard; protein; 579 AA.
 AC ADA28536;
 XX 20-NOV-2003 (first entry)
 DT Recombinant human lung tumour protein U523S #1.
 DE Cancer; lung cancer; gene therapy; vaccine; human;
 XX lung squamous cell carcinoma.
 KW
 XX

OS Homo sapiens.
XX
XX US2003064947-A1.
XX
XX 03-APR-2003.
PD
XX
XX 30-NOV-2001; 2001US-00007700.
PF
XX
XX 18-MAR-1998; 98US-00040802.
PR 27-JUL-1998; 98US-00123912.
PR 22-DEC-1998; 98US-00221107.
PR 02-APR-1999; 99US-00285479.
PR 17-DEC-1999; 99US-0046396.
PR 30-DEC-1999; 99US-00476496.
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
PR 04-APR-2000; 2000US-00542615.
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685696.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX
XX (CORI-) CORIXA CORP.
XX
XX Wang T, Wang A, Skeiky YAM, Li SX, Kalos MD, Henderson RA,
PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Matanabe Y, Peckham DW, Cai F, Foy TM;
XX
XX WPI; 2003-540798/51.
DR
XX
XX New isolated polynucleotides and polypeptides useful for diagnosing,
PT preventing and/or treating cancer, particularly lung cancer.
XX
XX
XX Claim 9; Page 285-287; 296pp; English.
PS
XX
XX The invention describes isolated polynucleotides and polypeptides useful
CC for diagnosing, preventing and/or treating cancer, particularly lung
CC cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
CC specification; complements of the nucleotide sequences cited above; at
CC least 10 contiguous residues of the nucleotide sequences cited above; a
CC sequence that hybridise to any of the nucleotide sequences cited above; a
CC stringent conditions; a sequence that is at least 75 or 90% identical to
CC the above nucleotide sequences; or degenerate variants of the above
CC nucleotide sequences. The composition and methods are useful in
CC diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC in gene therapy and in vaccines. This is the amino acid sequence of a
CC recombinant human lung tumour associated protein.
XX
XX
SQ Sequence 579 AA;
Alignment Scores:
Pred. No.: 5,47e-165 Length: 579
Score: 1875.00 Matches: 369
Percent Similarity: 78.70% Conservative: 78
Best Local Similarity: 64.96% Mismatches: 95
Query Match: 31.86% Indels: 26
DB: 7 Gaps: 6
US-09-270-437D-8 (1-3283) x ADA28536 (1-579)
QY 73 ATGAACAAGCTTTACATCGGAAACCTGAGCCCGCGCTCACCGCCGACGACTCCGGCAG 132
DB 1 MetAsnIleuTyrlleGlyAsnIleuSerGluAsnAlaIalProSerAspIleuGlnSer 20
QY 133 CTCCTTGGGACAGCAAGAGCTGCCCTGGCGGACAGGCTCTGTGAAGTCCGGCTACGCC 192
DB 21 IlePheIysAspAlaIalysrIleProValSerGlyProPheIeuValIysThrGlyTyrAla 40

QY 193 TTCGTGACTACCCCGACAGAACTGGGCGCATCCGGCCATCGAACCCCTCGGGTAA 252
DB 41 PheValAspCysProAspIleuSerTrpAlaIleGluAlaIleuSerGlyIys 60
QY 253 GTGAATTCAGTGGGAAATATCATGAAATGATTTACTAGTCTCTAAAGAGAGC 312
DB 61 IleGluIeuHsIsgIlySerProIleGluValGluHsSerValProIlyAsrGlnArgIle 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCACCTGAGTGGAGGCTGTGATGAGCACT 372
DB 81 ArgIlySerGlnIleArgAsnIleProProHsIleGlnITTPGIuValIeuAspSerIeu 100
QY 373 TTGGCTCAATATGGACAGTGGAGAAATGTGAAACAAGTCAACACAGACACAGAAACCGCC 432
DB 101 IeuValGlnTyrgIyValValIalGluSerCysIguIguIValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCACGTCACATATGCAACAAGAAAGAAAGCAAAATATGCCATGGAGAAAGCTAAC 492
DB 121 ValValAsnValThrIySerSerIysAspGlnAlaArgGlnAlaIeuAspIlySerAsn 140
QY 493 GGGCATCAGTTTGAAGAACTACTCTTCAAGATTTCTCAATCCCGATGAGAGAGTGGAC 552
DB 141 GlyPheGlnIleuGluAsnPheThrIeuIySerValAlaIatyrIleProAspGluThrAla 160
QY 553 TCCCTTGGCCCTCAG-----CGAGCCAGCGT-----GGGACCACTTCTCCCG 600
DB 161 GlnGlnAsnProIeuGlnGlnProArgGlyIyArgArgIyIeuGlyGlnArgIySerSer 180
QY 601 GAGCAAGGCGACGCCCCCTGGGGGCACTTCAGAGCGACACAGATTTATTCCTGGTGGG 660
DB 181 ArgIlyGly---SerProGlySerValSerIysGlnIysProCysAspIeuProIeuArg 199
QY 661 ATCTGTGTCGCCACCCAGTTTGTGTGTCATCCGAAAGAGAGGCGTGCCTTAAG 720
DB 200 IeuIeuValProThrGlnPheValGlyAlaIleIleGlyIySgIyGlyAlaThrIleArg 219
QY 721 AACATCACTAAGCAGACCCAGTCCCGGATGATATCCATAGAAAGAAAGAACTCGAGCT 780
DB 220 AsnIleThrIySgIlnThrGlnSerIyIleAspValHsArgIySgIyAsnAlaGlyAla 239
QY 781 GCAGAGAAAGCTGTCACCATTCACAGCCACCCCAAGGGGACTTGTGAAGCATGCCGATG 840
DB 240 AlaGlnIySerIleThrIleIeuSerThrProIuGlyIyThrSerAlaIaCylIySer 259
QY 841 ATTCTTGAATTCATGCAAGAAAGAGAGAGAGATGACCAAACTAGCCGAAGATTCCTGTG 900
DB 260 IleIeuGluIleMetHsIySgIyAlaGlnAspIleIySerIleThrGlnIuIleProIeu 279
QY 901 AAAATCTTGGCACACAATGCTGTGTGGAAGACTGATTTGAAAAGAAAGAGCAGAAATTTG 960
DB 280 IyIleIeuAlaHsIAsnAsnPheValGlyArgIeuIleGlyIySgIyGlnIyArgAsnIeu 299
QY 961 AAGAAATTCAGATGAAAGAGAGCAAGCAATATCAATTCATCTTTCAGATTTGAGC 1020
DB 300 IySylIyIleGluGlnAspThrAspThrIySylIeThrIleSerProIeuGlnIuIleThr 319
QY 1021 ATATACAAACCCGGAAGAAACCATCATCTGTGAAGGGACAGTGTAGGCTGGCCAGTGT 1080
DB 320 IeuTyAsnProIeuGlnIyThrIleThrValIySgIyAsnValGluIyThrCysAlaIySAla 339
QY 1081 GAGATTCAGATTTATGAAAGAGTGCCTGAGGCGCTTTGAAAATGATATGCTGCTGTAAAC 1140
DB 340 GluIguIguIleMetIySylIeArgIySerIyGluIyAsnAspIleAlaIerMetAsn 359
QY 1141 ACCCATCC-----GATATCTTC----- 1158
DB 360 IeuGlnAlaHsIleuIleProGlyIeuAsnIeuAsnAlaIeuGlyIeuPheProIyThr 379
QY 1159 -----TCAGCCTGTACCCCATCACACAGATTGAGCTGGCCCGCAT 1200
DB 380 SerGlyMetProProIyThrIySerGlyProProIySerAlaMetThrProIyIyProIyGln 399
QY 1201 CATCACTCTTATCCAGAGAGAGATTTGAAATCTCTCATCCCAACCAAGGCTGTGGCG 1260

```
Db      400 PheGluGln---SergIuThrGluThrValHisLeuPheIleProAlaLeuSerValGly 418
QY      1261 GCCATCATCGGGAGAGAGGGGACACATCAACAGCTGGCGAGATTGCCGGAGCTCT 1320
Db      419 AlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSer 438
QY      1321 ATCAAGATTGCCCTCGCGAGAGGCCAGACGTCAGCGAAAGATGGTCATCATCAACCGGG 1380
Db      439 IleLysIleAlaProAlaGlnAlaProAspAlaLysValArgMetValIleIleThrGly 458
QY      1381 CCACCGGAGAGCCCGAGTTCAGGGAGCGGATCTTTGGGAAACTGAAAGGAAAG 1440
Db      459 ProProGlnAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLysGlnGlnSer 478
QY      1441 TTCTTTAACCCCAAGAGAGAGTGAAGCTGAGAGCGCATATCAGAGTGCCCTCTTCACA 1500
Db      479 PheValSerProLysGlnGlnValLysLeuGlnAlaHisIleArgValProSerPheAla 498
QY      1501 GCTGGCCCGGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAACCTGCAAGACTTAACGAT 1560
Db      499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGlnLeuGlnAsnLeuSerSer 518
QY      1561 GCAGAGATCATGCTGCTGCTGACCAAGCCAGCATGAAATGAGGAAGTGATCGTCAGA 1620
Db      519 AlaGlnValValAlaProArgAspGlnThrProAspGlnAsnAspGlnValValLys 538
QY      1621 ATTATCGGGCACTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGAAATGTACAA 1680
Db      539 IleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlnIleLeuThr 558
QY      1681 CAGGTGAGCAGCAGCAGCAGAAA 1704
Db      559 GlnValLysGlnHisGlnGlnGln 566
```

Search completed: July 23, 2004, 11:05:23
Job time : 180.09 secs

Result	Query				Description	
No.	Score	Match	Length	DB	ID	
1	1934	32.9	577	3	US-09-661-855-2	Sequence 2, Appl1
2	1875	31.9	579	4	US-09-643-597-348	Sequence 348, Appl1
3	1875	31.9	579	4	US-09-542-615A-348	Sequence 348, Appl1
4	1875	31.9	579	4	US-09-606-621B-348	Sequence 348, Appl1
5	1868	31.7	579	4	US-09-643-597-176	Sequence 176, Appl1
6	1868	31.7	579	4	US-09-480-884A-176	Sequence 176, Appl1
7	1868	31.7	579	4	US-09-642-615A-176	Sequence 176, Appl1
8	1868	31.7	579	4	US-09-606-621B-176	Sequence 176, Appl1
9	263.5	4.5	644	1	US-08-021-608D-2	Sequence 2, Appl1
10	263.5	4.5	644	1	US-08-726-160-2	Sequence 2, Appl1
11	263.5	4.5	644	5	PCT-US94-01789-2	Sequence 2, Appl1
12	262	4.5	643	1	US-08-021-608D-10	Sequence 10, Appl1

13	262	4.5	643	1	US-08-726-160-10	Sequence 10, App1
14	262	4.5	643	1	PCT-US94-01782-10	Sequence 10, App1
15	255	4.3	590	1	US-08-021-608D-8	Sequence 8, App1
16	255	4.3	590	1	US-08-126-160-8	Sequence 8, App1
17	255	4.3	590	5	PCT-US94-01782-8	Sequence 8, App1
18	215	3.7	48	3	US-09-261-855-20	Sequence 20, App1
19	210	3.6	48	3	US-09-261-855-24	Sequence 20, App1
20	210	3.6	49	3	US-09-261-855-18	Sequence 18, App1
21	207	3.5	530	1	US-08-187-793-4	Sequence 18, App1
22	201	3.4	47	3	US-09-261-855-21	Sequence 21, App1
23	200	3.4	49	3	US-09-261-855-22	Sequence 21, App1
24	197	3.3	47	3	US-09-261-855-17	Sequence 17, App1
25	196.5	3.3	720	4	US-09-252-991A-10792	Sequence 12, App1
26	190.5	3.2	705	4	US-09-252-991A-10792	Sequence 11, App1
27	179	3.0	657	4	US-09-252-991A-24009	Sequence 21, App1
28	178.5	3.0	657	4	US-09-252-991A-18001	Sequence 21, App1
29	177	3.0	47	3	US-09-261-855-23	Sequence 23, App1
30	176	3.0	47	3	US-09-261-855-19	Sequence 19, App1
31	176	3.0	742	4	US-09-252-991A-12659	Sequence 12, App1
32	173.5	2.9	1706	4	US-09-252-991A-1760	Sequence 12, App1
33	170.5	2.9	333	1	US-08-187-793-2	Sequence 2, App1
34	169	2.9	1476	4	US-09-252-991A-2947	Sequence 29, App1
35	168.5	2.9	745	4	US-09-252-991A-29470	Sequence 29, App1
36	168.5	2.9	748	4	US-09-252-991A-18427	Sequence 18, App1
37	167.5	2.8	1228	4	US-09-252-991A-17764	Sequence 17, App1
38	167	2.8	987	4	US-09-252-991A-27483	Sequence 27, App1
39	166	2.8	783	4	US-09-252-991A-19431	Sequence 19, App1
40	166	2.8	871	4	US-09-252-991A-18035	Sequence 18, App1
41	164	2.8	654	4	US-09-252-991A-25423	Sequence 25, App1
42	164	2.8	830	4	US-09-252-991A-27142	Sequence 27, App1
43	163	2.8	685	4	US-09-252-991A-13184	Sequence 13, App1
44	162.5	2.8	955	4	US-09-252-991A-24254	Sequence 24, App1
45	162	2.8	681	4	US-09-252-991A-24267	Sequence 24, App1

ALIGNMENTS

```

RESULT 1
US-09-261-855-2
; Sequence 2, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ. ID NOS.: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ. ID NO. 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-2

```

Alignment Scores:

Pred. No.:	9.92e-163	Length:	577
Score:	1934.00	Matches:	381
Percent Similarity:	79.10%	Conservative:	77
Best Local Similarity:	65.80%	Mismatches:	95
Query Match:	3.86%	Indels:	26
DB:	3	Gaps:	5

[illegible]

```

QY 193 TTGGTGAAGTACCCCGACAGAACTGGGCGATCCGGCCATCGAGACCTCTCGGGTAA 252
Db PheValAspCysProAspGluHisThrPheMetLysAlaIleGluThrPheSerGlyLys 60
QY 253 GTGGAATTGCAAGGAAATATGATGGAAGTATTAATCAAGTCTTAAAGCTAAAGAGC 312
Db ValGluLeuGlnGlyLysArgLeuGlnMetGlnHisSerValProLysLysGlnArgSer 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCATCTGACGTGGAGGTGTGGATGACTT 372
Db ArgLysIleGlnIleArgSerHisLeuProGlnLeuArgThrGlnValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGGAACAGTGAAGATGTGAACAGTCAACAGACAGACAGAAACCCGC 432
Db LeuAlaGlnGlyLysThrValGlnLysCysGlnGlnValAsnThrGlnSerGlnThrAla 120
QY 433 GTTGTCAAGTCATATGCAACAGACAGAAAGAAATAGCATGAGAGAGTAAAGC 492
Db ValValAsnValThrLysSerAsnArgGlnGlnThrArgGlnAlaIleMetLysLeuAsn 140
QY 493 GGGCATCAGTTGAGAACTACTCTTCAAGATTCTCATCTCCGATGAGAGAGTGAAC 552
Db GlnHisGlnLeuGlnHisAlaLeuLysValSerThrLeuProAspGlnGlnIleThr 160
QY 553 TCCCTTCGCGCCCTCAGCAGAGCCAGCGTGGGAGCACTCTCCCGGAGAGAGCCAC 612
Db 161 -----GlnGlyProGlnAsnGlyArgArgGlyGlyPheGlySerArgLysGlnProArg 178
QY 613 -----GCCCTGGGGGACCTCTCAGGCGACAGACAGATGATTTCCGCTG 657
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspLeu 198
QY 658 CGGATCGGTGTCGCCACCCAGTTGTGTGGTCATCATCGGAAAGAGGCTTGACCA 717
Db 199 ArgLeuLeuValProThrGlnGlyValGlyAlaIleIleGlyLysGlnGlyAlaThrIle 218
QY 718 AAGACATCATCTAGACAGACCCCGGCTGATGATCCATGAAAGAAAGAAAGCTGGA 777
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGlnAsnAlaGly 238
QY 778 GGTGAGAGAGCGCTGACATCCATGACACCCCGAGAGGAGGACTTCTGAGAGTCCGC 837
Db 838 ATGATCTTGAATCATGACAGAAAGAGGAGATGAGACCAACTAGCCGAGAGATTCCT 897
Db 259 MetIleLeuGlnIleMetHisLysGlnAlaLysAspThrLysThrAlaAspGluValPro 278
QY 898 CTGAAATCTTGGCACACATGGCTGTGGAAGACTGATTTGGAAGAAAGAGAGCAAAAT 957
Db 279 LeuLysIleLeuAlaHisAsnAspPheValGlyArgLeuIleGlyLysGlnGlyArgAsn 298
QY 958 TTGAGAGAAATGTAACATGAAACAGGAGCAAGTAATCAATCTCATCTTTCAGAGATTG 1017
Db 299 LeuLysLysValGlnGlnAspThrGlnThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 1018 AGCATATTAACCCCGAAAGAACATCATCTGTGAAGGACACATGAGGCTGTGCCAGT 1077
Db 319 ThrLeuLysAsnProGlnArgThrIleThrValLysGlyAlaIleGlnLysCysArg 338
QY 1078 GCTGAGATTAAGATTATGAAGAACTGCTGAGAGCTTTGAAATGATATGCTGCTGT 1137
Db 339 AlaGlnGlnGlnIleMetLysValArgGlnAlaLysArgGlnAsnAspValAlaAlaMet 358
QY 1138 AAGACCACTCC-----GGATCTTC----- 1158
Db 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAla 378
QY 1159 ---TCCAGCTGTACCCCGATCACAG-----TTTGGCCGCTTCCCGAT 1200
Db 379 SerSerSerAlaValProProProProSerSerValThrGlyAlaAlaProLysSerSer 398

```

```

QY 1201 CATCACTTTATCCAGACAGAGATTGTGATCTCTTATCCCAACCCAGGCTGTGGC 1260
Db 399 PheMetGlnAlaProGlnGlnGlnMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1261 GCCATCATCGGAGAAAGGGGGGCACATCAAAACAGCTGGGAGATTGGCCGAGCCTCT 1320
Db 419 AlaIleIleGlyLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer 438
QY 1321 ATCAAGATTGCCCCCTGCGAAAGGCCAGACGTCAGCGAAAGAGATGTCAATCAACCGGG 1380
Db 439 IleLysIleAlaProProGlnThrProAspSerLysValArgMetValValIleThrGly 458
QY 1381 CCACCGAAGCCCACTTCAAGGCCAGGAGCGATCTTTGGGAAACTGAAAGAGGAAAC 1440
Db 459 ProProGlnAlaGlnPheLysAlaGlnGlyArgLysGlyLysLeuLysGlnGln 478
QY 1441 TTCTTTAACCCTCAAGAGAGTGAAGTGAAGCGCATATCAAGTGCCTCTTCCACA 1500
Db 479 PhePheGlyProLysGlnGlnValLysLeuGlnThrHisIleArgValProAlaSerAla 498
QY 1501 GCTGCGCGGGTGATTTGGCAAAAGTGGCAAGACCGTGAACGAATCGACAACTTAACAGT 1560
Db 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGlnLeuGlnAsnLeuThrAla 518
QY 1561 GCGAAGTCAATGTCCTCTGTCACCAACGCCAGATGAAATGAGGAAAGTGAATCGTCA 1620
Db 519 AlaGlnValAlaValProArgAspGlnThrProAspGlnAsnAspGlnValIleValLys 538
QY 1621 ATTATCGGGCACTTCTTGTGTCAGACTGACACGCGAAGTCAAGGAAATTTGTACA 1680
Db 539 IleIleGlyHisPheLysThrAlaSerGlnMetAlaGlnArgLysIleArgSerIleLeuAla 558
QY 1681 CAGGTGAAGACAGAGAGAGCAAAATACCTCAGGAGTGCCTCAGAGCGCAGCAAG 1737
Db 559 GlnValLysGlnGlnHisGlnLysGlyLysSerAsnLeuAlaGlnAlaArgArgLys 577

```

RESULT 2

```

US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348

```

```

Alignment Scores: 1.7e-157 Length: 579
Pred. No.: 1875.00 Matches: 369
Score: 1875.00 Matches: 369
Percent Similarity: 78.70% Conserved: 78
Best Local Similarity: 64.96% Mismatches: 95
Query Match: 31.86% Indels: 26
DB: 4 Gaps: 6

```


US-09-270-437D-8 (1-3283) x US-09-643-597-348 (1-579)

```

QY 73 ATGAAGAGCTTTAATCGGGAACCTGAGCCCGCTCACGCCGACGACTCCGGCAG 132
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGlnAsnAlaIaProSerAspLeuGlySer 20
QY 133 CTCTTTGGGACAGGAAGCTGCCCTGGCGGACAGGTCTCTGTAAGTCCGGCTAGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TTCGTGACTACCCCGACGAGAACTGGCCATCCGGCCATCGAGACCTCTCGGGTAA 252
Db 41 PheValAspCysProAspGlnSerIlePheValLeuLysAlaIleGlnAlaLeuSerGlyLys 60
QY 253 GTGGAATTGCATGGGAAAAATCAATGGAAGTTGATTACTCAAGTCTCTAAAAGCTAGAC 312
Db 61 IleGlnLeuHisGlyLysProIleGlnValGlnHisSerValProLysAspGlnThrGly 80
QY 313 AGGAAATTCAGATTGGAATTCCTCTCTCACTCGACGTGGAGAGGTGTGATGACTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProPheIleGlnIleProGlnValLeuAspSerLeu 100
QY 373 TTGGCTCATATATGGACAGTGGAGAAATGTGGAACAAGTAAACAACAAGAACCCGC 432
Db 101 LeuValGlnTyrGlyValValGlnSerCysGlnGlnValAsnThrAspSerGlnThrAla 120
QY 433 GTTGCAACGTCACTATATGCAACAAGAGAGGAAAAATATGCAATGGAGAGGCTAAC 492
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCGATTGAGAACTACTCTCTCAAGATTCTCAATCCCGATGAGAGGTGAGC 552
Db 141 GlyPheGlnLeuGlnAsnPheThrLeuLysValAlaTyrIleProAspGlnThrAlaAla 160
QY 553 TCCCTTGCGCCCTCAG-----CGAGCCAGCGT-----GGGACCACTCTTCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlnArgGlySerSer 180
QY 601 GAGCAAGGCCAGGCCCTGGGGGCACTTCTCAGGCGACAGATTTGATTTCCCGTGGC 660
Db 181 ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGTGCTCCCAACCCAGTTTGTGTGTCATCATCGGAAAGAGGGCTTGACCATAA 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGlnGlyAlaThrIleArg 219
QY 721 AACATCTAGAGAGCCAGTCCCGGGTACATATCCATAGAAAAGAACTGTGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGlnAsnAlaGlyAla 239
QY 781 GCAAGAGAGCCTGTCAACATCCATCCAGCCCAAGGGGACTTGTGAAGCATGCCGATG 840
Db 240 AlaGlnLysSerIleThrIleLeuSerThrProGlnGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAATCATGCAAGAAAGAGCATGAGAACCAACTAGCCGAGAGATTCCTGTG 900
Db 260 IleLeuGlnIleLeuMetHisLysGlnAlaGlnAspIleLysPheThrGlnGlnIleProLeu 279
QY 901 AAATCTTGCCACACAATGGCTGTGTGGAAGACTGATTGAAAAGAGGCAAGAAATTG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlnGlyArgAsnLeu 299
QY 961 AAGAAATTCAGATGAAACAGGAGCAAGATTAACAATCATCTTTGGAGATTGAGC 1020
Db 300 LysLysIleIleGlnGlnAspThrAspThrLysIleThrIleSerProLeuGlnGlnLeuThr 319
QY 1021 ATTTAACAACCCGAGAAAGCAATCACTGTGAGAGGACAGATTGAGGCTGTGCAAGTCT 1080
Db 320 LeuTyrAsnProGlnArgThrIleThrValLysGlyAsnValGlnThrCysAlaLysAla 339
QY 1081 GAGATGAGATTATGAAAGAGCTGCGTGGAGGCTTTGAAATATATATCTGCTGTTAAC 1140
Db 340 GlnGlnGlnIleLeuTyrLysIleArgGlnSerTyrGlnAsnAspIleAlaSerMetAsn 359

```

```

QY 1141 ACCCACTCC-----GGATACCTC----- 1158
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuLysLeuPheProThr 379
QY 1159 -----TCCAGCTGTACCCCCATACCAAGTTTGCCCGTCCGGCAT 1200
Db 380 SerGlyMetProProProThrSerGlyProProSerAlaMetThrProProTyrProGln 399
QY 1201 CATCACTCTTATTCAGAGCAGAGATTGTGAATCTCTATCCCAACCAAGGCTGTGGCC 1260
Db 400 PheGlnGln--SerGlnThrGlnThrValHisLeuPheIleProAlaLeuSerValGly 418
QY 1261 GCCATCATCGGAGAGAGGAGGACACATCAACAGCTGGCGAGATTCCCGGAGCCTCT 1320
Db 419 AlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSer 438
QY 1321 ATCAAGATTGGCCCGGAGAGGCCCAAGCTCAGCGAAAGATGTGCTCATCACCGGG 1380
Db 439 IleLysIleAlaProAlaGlnAlaProAspAlaLysValArgMetValIleIleThrGly 458
QY 1381 CCACCGAAGCCCAAGTTCAGAGGCCAGAGGATCTTTGGAGAACTGAAAGAGAAAC 1440
Db 459 ProProGlnValGlnPheLysAlaGlnGlyArgGlyTyrGlyLysIleLysGlnGlnAsn 478
QY 1441 TTCTTTAAACCCCAAGAAAGATGAAGCTGGAAGGCAATTCAGAGTCCCTCTCCACA 1500
Db 479 PheValSerProLysGlnGlnValLysLeuGlnAlaHisIleArgValProSerPheAla 498
QY 1501 GCTGGCGGGGTGATTGGGAAAGTGGCAAGACCGTGAACGAACTGCAGAACTTAACAGT 1560
Db 499 AlaGlyAlaGValIleGlyLysGlyGlyLysThrValaGlnGlnLeuGlnAsnLeuSerSer 518
QY 1561 GCAGAAGTCATCGTGCCTCGTACCAAGCGCAGATGAAATGAGAGTATCGTCAGA 1620
Db 519 AlaGlnValValValProArgAspGlnThrProAspGlnAsnAspGlnValValLys 538
QY 1621 ATTAATCGGAGCTTCTTTGTAGCCAGACTGCAGAGCGCAAGATCAGGAAATTGTACA 1680
Db 539 IleThrGlnHisPheTyrAlaCysGlnValAlaGlnIleArgLysIleGlnGlnIleLeuThr 558
QY 1681 CAGGTGAAGCAGCAGAGCAGAAA 1704
Db 559 GlnValLysGlnHisGlnGlnGln 566

RESULT 3
US-09-542-615A-348
? Sequence 348, Application US/09542615A
? Patent No. 6518256
? GENERAL INFORMATION:
? APPLICANT: Wang, Tongtong
? APPLICANT: Fan, Liqun
? APPLICANT: Kalos, Michael D.
? APPLICANT: Bangur, Chaitanya S.
? APPLICANT: Hosken, Nancy A.
? APPLICANT: Fanger, Gary R.
? TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
? FILE REFERENCE: 210121.455C8
? CURRENT APPLICATION NUMBER: US/09/542.615A
? CURRENT FILING DATE: 2000-04-14
? NUMBER OF SEQ ID NOS: 350
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 348
? LENGTH: 579
? TYPE: PRT
? ORGANISM: Homo sapiens
US-09-542-615A-348

Alignment Scores:
Pred. No.: 1,76-157 Length: 579
Score: 1875.00 Matches: 369
Percent Similarity: 78.708 Conservative: 78

```

Best Local Similarity: 64.96% Mismatches: 95
 Query Match: 31.86% Indels: 26
 DB: 4 Gaps: 6

US-09-270-437d-8 (1-3283) x US-09-542-615A-348 (1-579)

```

QY 73 ATGAAACAAGCTTTACATCGGGAACCTGAGCCCGCCGTCACCGCCGACGACCTCGGAG 132
Db 1 Metasnlrleuylrilegylasmlseuerglunslalaalaproseraspheugluser 20
QY 133 CTCCTTGGGACAGAGAACTGCCCCCTGGGAGACAGTCTCTGCTGAGTCCGGCTACGCC 192
Db 21 Ilephylsaspalalyllleprovalserglpropheleuvallythrghlytzyala 40
QY 193 TTCTGAGACTACCCGACCCAGAACTGGCCCATCCGCCCTTCGAGACCCCTCCGGGTAA 252
Db 41 PheValaspCysProaspgluserTrrpalaleuylsalalleglualaleuSERglYlys 60
QY 253 GTGAAATTGCACTGGGAAATATCATGAGTTGATTCAGTCTCTTAAAGCTAAGGAGC 312
Db 61 lilegluleuhsaglYysProilegluvalgluhservalProlyasparglnarglle 80
QY 313 AGGAAATTCAGATTGCAAAACATCCCTCCTCACTGACGTGGAGGCTTGGATGACTT 372
Db 81 Argylsleuglnllearghsnlleproproh:slauglntrpgluvalleuaspserleu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGAAACAGTCAACAGACAGAAACCGCC 432
Db 101 leuvalglntyrghlyalvalgluserCysgluInvalasnthrasperglunthrala 120
QY 433 GTTGTCAAGTCACATATGCAACAAGAGAAAGAAATAATAGCCATGAGAGAGTAAAC 492
Db 121 ValValasValThrYrsererlyrsabpghnlalarglnlaleuaspysleuasn 140
QY 493 GGGCATCAGTTTGAAGACTACTCCTTCAAGATTCTCAATCCCGATGAAGGCTGAGC 552
Db 141 GlyphelgluenglnunshenPheThrleuylsValalalyrlleproaspglunthrala 160
QY 553 TCCCTTCGCCCCCTGAG-----CGAGCCAGCGT-----GGGAGCACTTTCGCCG 600
Db 161 GlnGlunshenProleuglnInProargghlyargarglyleuylglInarglyseSer 180
QY 601 GAGCAAGCCAGCCGCCCTGGGGGACTTTCAGGCGCAGACAGATGATTTCCCGCTGCG 660
Db 181 ArgghlIngly--serProglYserValSerlysglnlyProCysaspheuproleuarg 199
QY 661 ATCTGGTCCCAACCCAGTTTGTGTGTCATATCGGAAAGAGGCTTGACCATTAAG 720
Db 200 leuLeuValProthrghlnPheValglYalallellysglnghlyalathrllearg 219
QY 721 AACATCACTAAGACACCCAGTCCCGGCTAGATTCATGAGAAAGAAAGAACTCGAGCT 780
Db 220 AsnleThrysglnthrglnserlyseilaspValhlsarglysglnasnlalaglyala 239
QY 781 GCGAGAGAGGCTGTCACATCATGACCCAGGAGGAGGACTTCTGAAGCATGCGGATG 840
Db 841 ATTCTGAAATCATGAGAAAGAGGAGGAGTAAACAATCATCTCTTTCAGATTGAGC 900
QY 901 AAAATCTTGGACACATGCTTGGTGGAGACAGTATGGAAAGAAAGAGGAAATTTG 960
Db 280 LyslleleuallahlsasasnsnphavalglYargleuilegylYysgluYlYarghsneu 299
QY 961 AAGAAATTTGAACATGAAACAGGAGCAAGATTAACAATCATCTCTTTCAGATTGAGC 1020
Db 300 LysylsllleglunghlnshpThrasphThryslleThrlleserproleuglnleuThr 319
QY 1021 ATATACAAACCGGAAAGAACATCACTGTGAAGGGGACAGTTGAGGCTTGCCAGTCT 1080
Db 320 leuYrshenProglunghThrlleThrVallysglnValgluThrCysAlalyala 339

```

```

QY 1081 GAGATAGAGATTATGAAGAAAGCTGCGTGGAGCCTTTGAAAATGATATGCTGCTTAAAC 1140
Db 340 GlnuInghlulemetCyslysllearggluserYrglnasnspllealeasermetasn 359
QY 1141 ACCCACTCC-----GGATACTTC----- 1158
Db 360 leuGlnlahlslleuileProglYleuasnleuasnlaaleuYleuPheProbrothr 379
QY 1159 -----TCCAGCCTGTACCCCATCAACAGTTTGGCCCGCTCCCGCAT 1200
Db 380 SerghlymetProproproThSerghlyProproseralameThnProProtyrProglIn 399
QY 1201 CATCACTCTTATCCAGAGACAGAGATTTGATCTCTTCATCCCAACCCAGGCTGGGC 1260
Db 400 PhegluIn-----SerghlunThrglnThrValhslleuPheleProaleuSERValghly 418
QY 1261 GCCATCATCGGAGAAAGAGGGGCAACATCAACAAACAGTGGGAGATTGCGCGGACCTCT 1320
Db 419 AlallelleghlyysghlnghlylnhsllelysglnseuSERargPhehlaaglyalaser 438
QY 1321 ATCAAGATTGCCCTCGGAGAGGCCAGACGTCAAGGAGAAAGATGCTCATCAACCGG 1380
Db 439 llelysllealelProalaglnlhaproaspalalyValargmetValillelthrgly 458
QY 1381 CCACCGGAAACCCAGTTTCAAGGCCCAAGAGGATCTTTGGGAAACTGAAAGAGAAAC 1440
Db 459 ProProglunlhalghlnPheylsalaglnghlyargllyrghlyrghlyrghlyrghly 478
QY 1441 TTCTTTAAACCCCAAGAAAGATGAAGCTGAAAGGCGCATATCAGATGCTCTTCCACA 1500
Db 479 PhevalserProlysergluInvallyseuGlnlahlslleargvalProserPheala 498
QY 1501 GCTGGCCGGGTATTTGGCAAGAGTGGCAAGACCTGTAAAGAACTGCAAGAACTTAACAG 1560
Db 499 AlaGlyargValilleghlyysghlyrghlyrghlyrghlyrghlyrghlyrghlyrghly 518
QY 1561 GCAGAAATCATGCTGCTGCTGTACCAAGCCAGATGAAGAAATGAGAGAGTATGTCAG 1620
Db 519 AlaGlnValValProargaspghlnthrProaspghlnshspghlnValVallyls 538
QY 1621 ATATATGGGCACTCTTTGCTAGCCAGACTGCAAGGCGCAAGATCAGGAAATTTGAC 1680
Db 539 lleThrglyshpshenThryalacyglInvalAlaglnhrgylsllleghlnluleuthr 558
QY 1681 CAGGTGAAGCAGCAGAGACAGAA 1704
Db 559 GlnVallysglnhslnglnghln 566

```

RESULT 4
 US-09-606-421B-348
 : Sequence 348, Application US/09606421B
 : Patent No. 6531315
 : GENERAL INFORMATION:
 : APPLICANT: Wang, Tonglong
 : APPLICANT: Fan, Liqun
 : APPLICANT: Kalos, Michael D.
 : APPLICANT: Bangur, Chaltanya S.
 : APPLICANT: Hosken, Nancy
 : APPLICANT: Fanger, Gary R.
 : APPLICANT: Li, Samuel X.
 : APPLICANT: Wang, Aijun
 : APPLICANT: Skeiky, Yassir A.W.
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 : TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 : FILE REFERENCE: 210121.455C9
 : CURRENT FILING DATE: 2000-06-28
 : NUMBER OF SEQ ID NOS: 358
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 348
 : LENGTH: 579
 : TYPE: PRT

ORGANISM: Homo sapiens
US-09-606-421B-348

Alignment Scores:
Pred. No.: 1,7e-157
Score: 1875.00
Percent Similarity: 78.70%
Best Local Similarity: 64.96%
Query Match: 31.86%
DB: 4
Length: 579
Matches: 369
Conservative: 78
Mismatches: 95
Indels: 26
Gaps: 6

US-09-270-437D-8 (1-3283) x US-09-606-421B-348 (1-579)

QY 73 ATGAACAGCTTTATCATCGGAACTGAGCCCGCGCTACCGCCGAGCCTCGGCGAG 132
Db 1 Metanlyslleutyllelyslneusergluasnalaalproserhsplneutlser 20
QY 133 CTCTTTGGGAGACAGAAAGCTGCGCCCTGGCGGAGAGGTCTGTGAAGTCCGCTACGC 192
Db 21 llephelysaspalalyserleprovalserglyprophellevalllyserthglytyrala 40
QY 193 TTCCGAGACTACCCCGACAGAACTGGGCGATCCCGCCATCGACACCTCTCGGGTAA 252
Db 41 PhevalaspCysproasppluserltpalaleuylsalalleglualaleuserglylys 60
QY 253 GTGAATTGCAATGGGAAATATCATGAGATTGATCTAGTCTTAAAGCTAAGAGAGC 312
Db 61 llegluleuhsiglylyleproilleglualgluhsservalprolysarglunrlyle 80
QY 313 AGGAAATTCAGATTTCGAAACATCCCTCTCACTGACGTGGAGGTGTGATGACTT 372
Db 81 Argylsleuglnlleargasnilleproprohlsleuglnltpgluvalleuaspserleu 100
QY 373 TTGCTCAATATGGGAGAGTGGAGATGTGAACAAGTCAACACAGACACGAAACCGCC 432
Db 101 leuvalglunrlyglvalvalglusercysglunlvalasntraspsergluthrala 120
QY 433 GTTGTCAACGTACATATGCAACAAGAGAAAGAAATAGCAATGACAGACTAAGC 492
Db 121 ValValasnValThrlyrserSerlyaspolnlaarglnalaleuaspLyseuasn 140
QY 493 GGGATATAGTTTGGAACTACTCTTCAAGATTTCATCATCCGCGATGAGAGGTGAGC 552
Db 141 Glyphelglneuglnuasnphetheleuylvalalalyrilleproaspgluthrala 160
QY 553 TCCCTTCGCGCCCTCGAG-----CGAGCCAGCGT-----GGGACCACTTCCCGG 600
Db 161 GlnglnasnProleuglnlnproarglyargllyleugllyglunrlyser 180
QY 601 GAGCAAGCCACGCCCTGGGGGCACTTCTCAGGCGCAGACAGATTGATTTCCCGTCCGG 660
Db 181 Argglngly---serprogllyserValserlyseglunrlyserocyaspleuProleuarg 199
QY 661 ATCTGTGCTCCCAACCGATTTGTGTGTCATCATCGAAAGAGGAGCTTACCATAAG 720
Db 200 leuileuvalProthrlnphevalgllyalaleillelyysglunrlyalathrillearg 219
QY 721 AACATCACTAAGAGAGCCAGTCCCGGTAGATATCATGAAAGAACTCTGAGACT 780
Db 220 Asnillethrlysglnthrlnserlylleaspvalhlsarglysglnuasnlaaglyala 239
QY 781 GCAGAGAGGCTGTCACTCATCCATGCAACCCGAGAGGAGCTTGAAGATCCCGCAG 840
Db 240 AlagluylserlethrilleuserThrproglnglythserlalaacylsysSer 259
QY 841 ATTCTTGAATCATGCAAGAGAGGAGATGAGACCAACTAGCCGAGATTCTCTCG 900
Db 260 lleleuglnullewethlsysglualaglnaspillelysbethehrlnuglnleproleu 279
QY 901 AAAATCTGGACACAAATGCTGTGTTGAAGCTGATTGAAAGAAAGAGGCAAAATTTG 960
Db 280 Lysilleuvalahlsasnasnphelvalgllyargleuilleglysglnugllyargasnleu 299

QY 961 AAGAAATTTGAATGAAACAGGACCAAGATPACATCTTTGACGATTTGAGC 1020
Db 300 LyslysllelglunlaspThrlysrlylethrilleSerProleuglnleuThr 319
QY 1021 ATATCAACCCGGAAGAAACCATCACTGTGAAGGCGACAGTGTGAGCCCTGTCCAGTGT 1080
Db 320 leuTyrsasnproglunrlythrillethrvallysglyasnvalglunrlysalalysala 339
QY 1081 GAGATAGAGATTATGAGAGAGCTGCGTGGAGCTTTGAAATGATGTGCGTGTAAAC 1140
Db 340 GlnglnuglnullewethlysbelysrilleargluserTyrgluasnaspilleasertasn 359
QY 1141 ACCCACTCC-----GATACTTC----- 1158
Db 360 leuglnalahlleuilleproglyleuasnleuasnalaaleuglyneupheprothr 379
QY 1159 -----TCCAGCTGTAAACCCCATCACCAAGTTGGCCGTTCCGCGAT 1200
Db 380 SerglyMetProProProthrSerllyProProSerlameThrProProTyProglin 399
QY 1201 CATCACTCTTATCCAGACAGAGATGTGATCTCTTCATCCCAACCGAGCTGTGGG 1260
Db 400 Pheglunl----serglunrlythrvalhlsleupheilleproalaleuservalgly 418
QY 1261 GCGATCATCGGAGAGAGGCGGACACATCAACAGCTGCGAGATTGCGCGAGCTCT 1320
Db 419 AlallellellysglnlyglnhlsilleysglnleuserArgPhalaglylaser 438
QY 1321 ATCAAGATTGCCCTCGGAGAGGCCCAAGCTCAGCGAAAGATGTGTATCATCACCG 1380
Db 439 llelysllealaproalaglnalaproaspalalyvalaJargmetvalillethrly 458
QY 1381 CCACCGAAGCCGATTCAGAGCCAGGAGCGATCTTTGGGAAATGAAAGAGGAAAC 1440
Db 459 Proproglunlagnlnpheylsalaglnlyargllylelygllysllelysglnunr 478
QY 1441 TTCTTAAACCCCAAGAAAGTGAAGCTGAGCGCATCATCAGAGTCCCTCTTCCACA 1500
Db 479 PhevalserProlysglnuglnvallyleuglnlahlslleargvalProserPhehla 498
QY 1501 GCTGCGCGGTGATTTGGCAAGGTGGCAAGACCGTGAACGATTCGACAACTTAAACAGT 1560
Db 499 Alaglyargvalillelygllysglylyllythrvalasnuglnleuasnleuser 518
QY 1561 GCAGAGTCACTCGGCTCTGACCAACCCGATGAAATGAGAAAGTATGTCGAGA 1620
Db 519 AlagluvalvalalProahgaspglnthrProaspplunsnspglunlvalVallys 538
QY 1621 ATTATCGGCACTTCTTGTAGCCAGACTGCAAGCGCAAGATCAGGAAATGTGACA 1680
Db 539 llethrcllyhsphetyralacysglnvalalaglnhargylsilleglnglnulleuThr 558
QY 1681 CAGGTAGAGCAGAGAGCAAGAA 1704
Db 559 Glunvalysglnhsglnngln 566

RESULT 5
US-09-643-597-176
Sequence 176, Application US/09643597
Patent No. 6426072
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Henderson, Robert A.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

```

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643.597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-643-597-176

Alignment Scores:
Pred. No.: 7,1e-157 Length: 579
Score: 1868.00 Matches: 368
Percent Similarity: 78.52% Conservative: 78
Best Local Similarity: 64.79% Mismatches: 96
Query Match: 31.74% Indels: 26
Gaps: 6

US-09-270-437D-8 (1-3283) x US-09-643-597-176 (1-579)

QY 73 ATGAACAAGCTTACATCGGAGCACTGAGCCCGCGGTACCCGCGAGCAGCTCCGCGAGC 132
Db 1 MetasnllyslentylrlllellyasnleusergluasnlaialaProserdspleugluser 20
QY 133 CTCTTGGGAGCAGAGAGCTGCCCCCTGGCGGAGCAGCTCTGCTGAAGTCCGGCTACGCC 192
Db 21 llelhelylaspralalyserlleprovalserglyProheleuvalylserthrglytryala 40
QY 193 TTGCTGAGTACCCCGCAGCAAGCTGGGCGCATCCGGCGCATCGAGCCCTCCGGGTAA 252
Db 41 PhevalAspCyseProspgluserlrralaleuylalaileluhlaaleuserglylys 60
QY 253 GTCGATTCGATGGGAAATCATGAGATGATCTCACTCTCTAAAAGCTAAGAGC 312
Db 61 llegluenuhlselylsProtlegluvalgluhlservalProlysrhrglnlrgyle 80
QY 313 AGGAAATTCAGATTGCAAAATCATCTCTCTACCTGCACTGAGAGTGTGGATGACTT 372
Db 81 ArglylsleuglnllelrgasnllleProkonhleleglntprgluvalleuAspserleu 100
QY 373 TTGCGCTAATATGGGACAGTGGAGATGTGGAACAAGTCAACAAGCAGACAGAAACGCC 432
Db 101 leuvalglntylrlyalvalgluserCysgluqluvalasnhrAspserglutprala 120
QY 433 GTTGTCAGGTCACATATGCAACAGAGAGAAATAGCCATGAGAACTAAGC 492
Db 121 ValValasnValthrlyrserlserlyAspqlnlaarglnlaaleuAspysleuasn 140
QY 493 GGGCATCATGTTGAGAACTACTCTTGAAGATTCTTACATCCCGGATGAAAGGTAGC 552
Db 141 GlYpHehlnleuglnuasnPhelrleuylvalAlaIatylrleProAspgrlmetAlaIa 160
QY 553 TCCCTCTCGCCCTCG-----CGAGCCAGCGT-----GGGACCACTCTTCCGG 600
Db 161 GlnqlnAsnProleuglnqlnProarglylArgrglleuglygluqlnarglylserSer 180
QY 601 GAGCAAGGACGAGCCCTGGGAGCACTTCTGAGGCGCAGACAGATGATTTCCGCTGGG 660
Db 181 Argqlnqlnly--SerProgllyserValserlysglnlyspCysAspLeuProleuArg 199
QY 661 ATCTGTGTCCTCCACCCAGTTGTGTGGTCCATCATCGGAAAGAGGGCTTGAACCTAAG 720
Db 200 leuLeuvalProthrlnPheValglYAlaIlellellylsgluqlnYAlaIthrIleArg 219
QY 721 AACATCACTAAGCAGACCACTCCCGGTAGATATCAATAAAAGAGAACTCGAGACT 780
Db 220 AsnlllethrlysglnthrglnserlylsleAspValhlsarglylsgluAsnAlaIyla 239
QY 781 GCAGAGAAGCCTGTACCATCATGACCAAGGAGGAGGAGCTTGAAGCATGCCGATG 840
Db 240 AlaqluylserllethrIleleuserthrProgluqlnYThrserAlaIaCylslyser 259

```

```

QY 841 ATTCTGAATATGCAAGAAAGCAGATGAGACCAACTAGCCGAGAGATTCCTCTG 900
Db 260 lleleuglnllethslslysglnlaIglAsprlleysPheThrsglnlleleu 279
QY 901 AAAATCTGGACACATATGCTGTGGTGAAGATGATGGAAGAAAGAGCAAAATTTG 960
Db 280 lyslleleuAlahlsAsnAsnPhelValglYargleuIlellylsgluqlnYArgAsnleu 299
QY 961 AAGAAATTTGAACATGAAACAGAGACCAAGATTAACAATTCATCTTTGACAGATTGAGC 1020
Db 300 lyslysllegluqlnAspThrAspThrlysllethrIleSerProleuglnleuThr 319
QY 1021 ATATACACCCGGAAGAACCATCATCTGTGAAGGCGACAGTTGAGGCTGTGCCAGTCT 1080
Db 320 leuTryAsnProgluArgThrIlethrVallysglnvalgluThrCysAlaIysAla 339
QY 1081 GAGATGAGATTATGAAAGACCTGGTGAGGCTTGAAGATGATGCTGCTGTAAAC 1140
Db 340 GluqluqlnleuMetlyslleArggluserlyrgluAsnAspIleAlaSerMetAsn 359
QY 1141 ACCCACTCC-----GGATCTTC----- 1158
Db 360 leuqlnlaIahlsleuIleProgllyleuAsnleuAsnlaIeuglyLeuPheProthr 379
QY 1159 -----TCCAGCTGTACCCCATACCAAGTTGGCCCGTCCGCGCAT 1200
Db 380 sergluMetProProthrserglyProProserAlaMetThrProProtyrProglin 399
QY 1201 CATCACTCTTATCCAGAGCAGAGAGATTTGGATTCCTTCATCCCAACCGAGCTGTGAC 1260
Db 400 Phegluqln---SerqluThrqluThrValhlsghlnPheIleProAlaIeuserAlaIy 418
QY 1261 GGCATCATCGGAGAGAGGAGGAGCAGACATCAACAGCTGGCAGATTCGCGGAGCTCT 1320
Db 419 AlaIlellellylsghlnlyghlnhlslelysglnleuserArgPheAlaIyAlaSer 438
QY 1321 ATCAAGATTCCTCTCGGAGAGGCCAGCTCAGCGAAAGATGGTCATCATCAACGGG 1380
Db 439 llelyleIAlaProAlaqluAlaProAspAlaIysValArgMetVallellethrly 458
QY 1381 CCACCGGAGCCCGAGTTCAAGGCCGAGGAGCGGATCTTGGGAAATGGAAGAGAAAC 1440
Db 459 ProProgluAlaqlnleuylsAlaqlnlyArgIlelyrlylysllelysglnuasn 478
QY 1441 TTCTTAACCCCAAGAGATGAGCTGGAAGCCCATATCGAGTCCCTCTTCCACA 1500
Db 479 PhevalserProlysglnuvallylsleuglnlaIahlslellyrValProserPheAla 498
QY 1501 GCTGGCCGGGTGATTGGCAAAAGGTGGCAAGACCTGTGAACGAACCTGCAAGCTTAACAGT 1560
Db 499 AlaqlYargValllellylsghlylylysthrValAsnqlleuqlnAsnleuserSer 518
QY 1561 GCAGAACTCATCTGCTCGCTCGTGAACCAAGCCAGATGAAATAGAGAAAGTATGTCAGA 1620
Db 519 AlaqluvalValProArgAspqlnThrProAspqluAsnAspqlnvalVallyls 538
QY 1621 ATATCGGAGCACTTCTTGTGACCACTGCAACAGCGCAAGATCAGGAAATGTACAA 1680
Db 539 llehrqllyshsPheYrAlaCysqlnvalAlaqlnArglysllelgluqlnleuThr 558
QY 1681 CAGGTGAAGCAGAGAGAGAA 1704
Db 559 Gluvallysglnhlsghlnqln 566

RESULT 6
US-09-480-884A-176
; Sequence 176, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongfong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.

```

; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C6
 ; CURRENT APPLICATION NUMBER: US/09/480,884A
 ; CURRENT FILING DATE: 2001-08-27
 ; NUMBER OF SEQ ID NOS: 330
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 176
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-480-884A-176

Alignment Scores:
 Pred. No.: 7,1e-157 Length: 579
 Score: 1868.00 Matches: 368
 Percent Similarity: 78.52% Conservative: 78
 Best Local Similarity: 64.79% Mismatches: 96
 Query Match: 31.74% Indels: 26
 Gaps: 6

US-09-270-437D-8 (1-3283) x US-09-480-884A-176 (1-579)

QY 73 ATGAAAGCTTTATCATGAGGAGACCTGAGCCCGCCCTCACCGCCGACGACCTCCGAGC 132
 Db 1 MetasulysleutyrlleglYAsnleuserglunsnlaialaProserAspleugluser 20
 QY 133 CTCTTTGGGGGACAGAAAGCTCCCTGGCGGAGACAGCTCTGCTGAAGTCGGGCTACGCC 192
 Db 21 lleheylsAepAlaAllysleleProvalserglYProphleuVallysthrGlYtYala 40
 QY 193 TTCTGAGTACACCCGACAGACTGGGCGCATCGCGGCATCGAGACCCCTTGGGTTAA 252
 Db 41 PheAlaspYspProhsBgluserlPrAlaleuylsAlalleglunaleuSerGlYlys 60
 QY 253 GTGGAATTGACAGGAAATCATGAAATGATTGATTACTACAGCTCTTAAAGCTAAGAGC 312
 Db 61 lleglunehlsiglYsProlleglunaleuSerValProlyshArglnArglle 80
 QY 313 AGGAAATTCAGATTCGAAACATCTCTCTCACTGAGTGGAGGTTGGATGAGCTT 372
 Db 81 ArgylsleuGlInleArgAsnleleProhlsleuGlInlPrGlValleuAspserleu 100
 QY 373 TTGGCTCAATATGAGGACAGTGGAGATGTGGAATCAATCAACAGACAGAAACGCC 432
 Db 101 leuValGlInlYrGlYValValgluserCysglunlValAsnthraspsergluthrAla 120
 QY 433 GTTGTCAAGCTCATATGACAGAGAGAAAGCAAAATAGCCATGAGAGAGCTAAGC 492
 Db 121 ValValAsnValthrYrserSerlyshAspGlInlAlarGlInlaleuAspLysleuAsn 140
 QY 493 GGGCATCATGTTGAGAACTAATCTCTTCAGATTTCTCAATCCCGATGAGAGAGTGC 552
 Db 141 GlyPheglInleuGlInsnPheThrleuYsValAlAltyrlleProAspGlInlMetAlaAla 160
 QY 553 TCCCTTCGGCCCTCAG-----CGAGCCCAAGCT-----GGGACCACTTCTCCCG 600
 Db 161 GlInlAsnProleuGlInlPrOArGlYArgArgglYleuGlYglInlArgglYserSer 180
 QY 601 GAGCAAGGCCACGCCCTGGGGGCACTTCTCAGCGACAGACAGATTTGTCGCGCGG 660
 Db 181 ArgGlInlY---serProglYserValserlyshlYserlYserCysAspLeuProleuAsn 199
 QY 661 ATCTGTGTCCCAACCAAGTTTGTGTGTCATCATCGAAAGGAGGCTTGAACATTAAG 720
 Db 200 leuValValProthrInlPheValAlglYAlAllelleglYlysGlInlYlathrleArg 219
 QY 721 AACATCATAGAGAGACAGCCAGTCCCGGATGATATCATAGAAAGAAAGAACTTGAAGCT 780
 Db 220 AsnlelthrYslelInlthrInlserlyshleAspValhlsArgylsGlInlAsnAlglYAla 239

QY 781 GCAGAGAGCCCTGTGACCATTCATGCCAACCCAGAGGGGACTTTGAGAGCTCCGAGT 840
 Db 240 AlaGlYsSerlethrleuSerThrProglInlYThrSerAlaAlaCyslysSer 259
 QY 841 ATCTTGAATCATCAGAAAGAGGACAGTGAACCAACTGACCGGAAAGATTCCTCTG 900
 Db 260 lleleuGlInlMetleHlsleGlInlAglInAspIlelYsPheThrGlInluleProleu 279
 QY 901 AAAATCTTGGACACAAATGGCTTGGTGGAAAGATGAAATGAAAGAGCAAAATTG 960
 Db 280 lyslleleuAlahlsAsnAsnPheValglYthrleuIleglYlysGlInlYrArgAsnleu 299
 QY 961 AAGAAATTTGAATGAAACAGGAGCCAAAGTAAATCATCTTTGACGAGATTGAGC 1020
 Db 300 lysYsllleglInlAspThrAspThrlyslleThrlyleSerProleuGlInluleuThr 319
 QY 1021 ATATCAACCCGGGAAAGAACCATCATCTGTGAAGGGGACAGTTGAGCCTGTGCCAGTGT 1080
 Db 320 leuYrAsnProglInlAspThrleThrVallyshlYAsnValgluthrCysAlAlYsAla 339
 QY 1081 GAGATGAGATTATGAAAGACCTGCGTGAAGCCCTTGAAGAAATGATATGCTGCTTTAA 1140
 Db 340 GlunlInluleMetYslelleArgluserlYrGlInlAsnAspIleAlSerMetAsn 359
 QY 1141 ACCCACTCC-----GGATCTTC----- 1158
 Db 360 leuGlInlAlhlsleulleProglYleuAsnleuAsnAlaleuGlYleuPheProthr 379
 QY 1159 -----TCCAGCCTGACCCCATACCAAGTTTGGCCCTTCCCGAT 1200
 Db 380 SerlYmetProProthrInlSerGlYProProserAlamethrProProlyrProhln 399
 QY 1201 CATCATCTTATCCAGAGACAGAGATTGTGATCTTTCATCCCAACCCAGAGCTGTGGC 1260
 Db 400 PheglunlIn---sergluthrInlthrValhlslelnhelleProAlaleuSerValAlY 418
 QY 1261 GCCATCATCGGAAAGAGGGGGACACATCAACAGCTGGGAGATTGCGCGAGCCTCT 1320
 Db 419 AlallelleglYysglInlYslnhlslelleuSerleuSerArgPheAlglYAlaser 438
 QY 1321 ATCAAGATTGCCCCCTGGGAGAGCCGAGAGGTCAAGCAAGATGTCATCATACCGGG 1380
 Db 439 lleYsllleAlProAlaGlInlAlProAspAlAlYsValhArgmetVallellethrGlY 458
 QY 1381 CCAACCGGAGCCAGTTCAAGGCCCAAGGACGAGATCTTGGGAACTGAAAGAGAAAC 1440
 Db 459 ProProglunlAglInPheYsAlaGlInlYArglleYrGlYslelYsleuGlInlAsn 478
 QY 1441 TTCTTAAACCCCAAGAGAGAGTGAAGCTGGAAGCCGATTCAGAGTGCCTTTCACA 1500
 Db 479 PheValserProlysglunlVallyshleuGlAlAlhlslelleArgValProserPheAla 498
 QY 1501 GCTGGCGCGGTGATTTGGCAAAAGGTGGCAAGACCGTGAACGAACTGCAGAACTTAACAGT 1560
 Db 499 AlaGlYrArgVallleglYysglYglYlYsThrValAsnleuIleuGlInlAsnleuSerSer 518
 QY 1561 GCAGAAAGTATCTGTGCTGTGACCAAGGCCAGATGAAATGAGAAAGATGCTGTGAGA 1620
 Db 519 AlaGlInlValValProArArgAspGlInlhrProAspGlInlAsnAspGlInlVallyls 538
 QY 1621 ATTATGGGCACTTCTTGTGACCAAGTGCACAGCTGCACAGCGCAAGATCAAGAAATGTACA 1680
 Db 539 llethrGlYhlsPheYrAlaCysGlInlAlAlaGlInlArgyllelleInluleuThr 558
 QY 1681 CAGGTGAGCAGAGAGCAGAA 1704
 Db 559 GlInlAlYsGlInlhlsglInlIn 566

RESULT 7
 US-09-542-615A-176
 ; Sequence 176, Application US/09542615A
 ; Patent No. 6518256
 ; GENERAL INFORMATION:

APPLICANT: Wang, Tongtong
 APPLICANT: Fan, Liqun
 APPLICANT: Kalos, Michael D.
 APPLICANT: Bangur, Chaelanya S.
 APPLICANT: Hosken, Nancy A.
 APPLICANT: Fanger, Gary R.
 TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 FILE REFERENCE: 210121.455C8
 CURRENT APPLICATION NUMBER: US/09/542.615A
 CURRENT FILING DATE: 2000-04-14
 NUMBER OF SEQ ID NOS: 350
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 176
 LENGTH: 579
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-542-615A-176

Alignment Scores:
 Pred. No.: 7.1e-157 Length: 579
 Score: 1868.00 Matches: 368
 Percent Similarity: 78.52% Conservative: 78
 Best Local Similarity: 64.79% Mismatches: 96
 Query Match: 31.74% Indels: 26
 Gaps: 6

US-09-270-437d-8 (1-3283) x US-09-542-615A-176 (1-579)
 QY 73 ATGACCAAGCTTTATCATGGGAAAGCTGAGCCCGCGTCAACCGCGAGCAAGCTCCGCGAG 132
 Db 1 MetAmlylsleutyllelglyasnleusergluasnlalalproserlsleuuser 20
 QY 133 CTTCTGGGAGAGAGAGCTGCCCTGGCGGAGAGCTGCTGTAAGTCCGCTTACGCGC 192
 Db 21 liepneulyspialalysileprovalsergilypropheluevallyshrgilytyala 40
 QY 193 TTGCTGACTACCCGACCAAGAACTGGGCAATCCGCGCAATCGAGACCCCTCGGGTAA 252
 Db 41 PhevalaspCysProaspjuserltpalaleuylsalaleglualaleuserglylys 60
 QY 253 GTCGATTCGATCGGAAATCATGAAAGTTGATTACTAGTCTCTAAAAAGCTAAGAGC 312
 Db 61 lieglleuhsileglysproillegluvalgluhiservalprolysharglnaglie 80
 QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCACCTGCACTGAGTGGAGGTGTGAGTGA 372
 Db 81 Arglylsleuqlnlelthargenilleproprohileuqlntrpgluvalleuaspserleu 100
 QY 373 TTGGCTCAATATGAGACAGTGAAGATGTGAAACAGTCAACACAGACAGAAACCGGC 432
 Db 101 leuvalglintyrglyvalalgluserCysgluglnvalasnthraspsergluthala 120
 QY 433 GTTGTCAGCTCATATGACACAGAGAGAGAAATAAGCCATGAGAGAGTAAAGC 492
 Db 121 Valvalasnvalthrlytserseulyaspelnlalarglnalaleuaplylsleuasn 140
 QY 493 GGCATCATGTTGAACTACTCTTCAGATTCTCAATTCCTCAATCCCGATGAAGAGTAA 552
 Db 141 Glphelelnleuqlnleuasnphethrleuylvalalalyrilleproaspglumetalala 160
 QY 553 TCCCTTCGCCCCCTCAG-----CGAGCCAGCGT-----GGGAGCACTCTCCCGG 600
 Db 161 GlnglnasnproleuglnglnlnproarglylArgrarglyleuqlnglnaglyserSer 180
 QY 601 GAGAGAGCCAGCGCCCTCGGGGACATTCAGAGCCAGACAGATGATTTCCTCCGCGG 660
 Db 181 Argnglyly--SerProglySerValserlysglnlysproCysapleuProleuarg 199
 QY 661 ATCTGATCCCAACCAAGTTGTGTGTCATCATCGAAAGAGGCTTGAACATAAG 720
 Db 200 leuvalvalprothrglnphevalglYalalletlelglysglnuglyalalthrlearg 219

QY 721 AACATCAATAGACAGACCCAGTCCGGGTAGANTCCATAGAAAAGAACTGTGAGCT 780
 Db 220 Asnlethrlysglnthrlnserlyleaspvalhsarglysgluasnlaleglyala 239
 QY 781 GCAGAGAGCTGACCATTCATCCAGCCCAAGGGGACCTTCGAAAGATGCGGATG 840
 Db 240 Alaglylsseerillethrleuuserlthprogluqlnthrseralalaclylyser 259
 QY 841 ATTCTGAAATCATGACAGAAAGAGCATGACCAAACTAGACCGAAAGATTCCTCTG 900
 Db 260 lieuglnleuethslslyglualaglnasplleuyspethrglnuglnleproleu 279
 QY 901 AAAATCTTGCAACAATGCTGTGTTGAAAGATGATGGAAGAAAGAGAGAAATTTG 960
 Db 280 LyslleuvalahlsasnasnphvalglYargyleuileglylysglnlylYargsmleu 299
 QY 961 AAGAAATTTGAACATGAAACAGGAGCAAGATTAACATCTCATCTTGTCAGATTGAGC 1020
 Db 300 LyslyslleuglnaspthraspthrlysllethrlyleserProleuglnleuethr 319
 QY 1021 ATATCAACCCGAAAGAACCATCATCTGTGAAGGGCAGATTGAGGCTGTGCCAGTCT 1080
 Db 320 leuYrznprogluargThrillethrvalysglYasnvalgluthrCyslalysala 339
 QY 1081 GAGATGAGATTATGAAGAGCTGCTGAGGCTTGAAGAAATGATATGCTGCTGTTAC 1140
 Db 340 Gluglnuglnleuethlysllearggluserlyrgluasnaspillelasermetasn 359
 QY 1141 ACCCACTCC-----GGATACTTC----- 1158
 Db 360 leuqlnlahlsleuilleproglyleuasnleuasnlaaleuqlyleuethrProthr 379
 QY 1159 -----TCCAGCTGTACCCCATCAACAGTTGGCCCGCTTCCGCGAT 1200
 Db 380 SerglymeProProProthrserglyProProseralameThrProProlyrProglin 399
 QY 1201 CATCACTTTATCCAGAGAGAGATTCGATCTTCATCCCAACCAAGCTGCGGCG 1260
 Db 400 Pheglugln---SerlthrglnthrValahlslnphelleprohaleuserValgly 418
 QY 1261 GCCATCATCGGAGAGAGGGGACACATCAAAACAGCTGCGAGATTCGCGGAGCTCT 1320
 Db 419 Alalletlelglysglnlelyglnhlslelysglnleuserlthpheelaglylaser 438
 QY 1321 ATCAAGATTGCCCTCGGAGAGCCCAAGAGCTGACGGAAGAGATGTCATCATACCGGG 1380
 Db 439 lieylsillealprohlaqlunlaProaspalalyvalargmetvalleillethrgly 458
 QY 1381 CCACCGGAAGCCCATGTTCAAGGCCCAAGAGCGGATCTTGGGAAACGAAAGAGAAAC 1440
 Db 459 ProProglualaglnphelysalaqlnglylYargyletlylYstlelysglnuglnasn 478
 QY 1441 TTCTTTAACCCCAAGAGAGATGAGCTGGAAGCGCATATCAAGATGCCCTCTTCA 1500
 Db 479 PhevalserProlysglnuglnvalysleuqlnlahlslelthargvalProserPheala 498
 QY 1501 GCTGCGCGGTTGATTGCAAAAGGTGCAACAGCTGAAACGATCGCAAGCTTAAACAGT 1560
 Db 499 AlaglylYargvalilleglylysglylYsthrValasnuglnleuqlnleuasnleuser 518
 QY 1561 GCGAAGATCATGCTGCTCTGTCACCAAGCCAGATGAAATGAGAAATGATGATGTCAGA 1620
 Db 519 AlagluvalvalalProargaspqlnthrprospulnshaspqlnvalValVallys 538
 QY 1621 ATTATGCGGACATCTTGTGTAAGCACTGCAAGCGCAAGATCAGAGAAATTTGACA 1680
 Db 539 liehrlyshlspehrylhalCysqlnvalalaglntrgylslelnglnlulleuethr 558
 QY 1681 CAGGTGAGCAGCAGAGCAGAAA 1704
 Db 559 glnvalysglnhsglnglngln 566

RESULT 8

US-09-606-421B-176
; Sequence 176, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-176

Alignment Scores:
Pred. No.: 7.1e-157 Length: 579
Score: 1868.00 Matches: 368
Percent Similarity: 78.52% Conservative: 78
Best Local Similarity: 64.79% Mismatches: 96
Query Match: 31.74% Indels: 26
Gaps: 6

US-09-270-437D-8 (1-3283) x US-09-606-421B-176 (1-579)

QY 73 ATGAACAAGCTTACATGGGAACTGAGCCCGCCGTCACCGCCGACGACCTCCGGCAG 132
Db 1 Metasnysleutyrlieglyasnleusergluasnlaalaproseraspneuqluser 20
QY 133 CTCTTTGGGACAGAGAGCTGCGCCCGGGACAGAGCTCTGCTGAAGTCCGGCTACGCC 192
Db 21 Ileheylsaspalalysilleprovalserdylpropheleuvalyssthglytyrala 40
QY 193 TTGCTGAGTACCCCGACAGAACTGGGGCATCGCGCCATCGAGACCTCTGGGTAA 252
Db 41 Phevalaspcysproaspjusertrpalauleuysalalleglualaleuserglylys 60
QY 253 GTGAATTGATGGGAAATATGAAAGTTGATTACTCACTCTCTAAAGCTAAGAGC 312
Db 61 Ilegluenuhsglyysproillegluvalglwhiservalprolysarglunargile 80
QY 313 AGGAATTCAGATTGAAACATCCCTCCCTCACCTGAGGGAGGCTTGATGACTT 372
Db 81 Arglylsleuglntleahgaenlleprohishleuindtrpgluvalleuaspsrleu 100
QY 373 TTGGCTCAATATGAGAGAGTGGAGAAATGTTGAAACAAGTCAACAAGAGAGAGTGG 432
Db 101 LeuvalaGlntyrclvalValaGlusercysgluGlntlnhraspsrsergluntrala 120
QY 433 GTTGTAAAGTCAATATGCAACAAGAGAGAAATAGCAATAGCAAGAGTAAAGC 492
Db 121 ValValaAsnValThrlyrserSerlysaSpGlnAlaArglnAlaleuaspysleuaen 140
QY 493 GGGCATCGTTGAGAACTACCTTCATGAGATTCTCTACATCCCGGATGAAGGTGAGC 552
Db 141 GlypheglntleuglunsnphetrlnrleuylsValAlaIatylleproaspjmetalaala 160
QY 553 TCCCTTCGCCCCCTCAG-----CGAGCCAGCGT-----GGGACCACTCTTCGG 600
Db 161 GlnGlnAsnProleuglntlnProaArglyArgArgglyleuclglylnarglyserSer 180
QY 601 GAGCAAGGCGACGCGCTGGGGGCACTTCTCAGGCGCAGACAGATTGATTTCCCGCTGCG 660

Db 181 ArgGlnGly---SerProIlySerValSerlysglnlyserProCysaspneuProleuArg 199
QY 661 ATCTGGTCCCAACCCAGTTTGTGGTCCATCATCGAAAGAGAGGCTTGACCAATAG 720
Db 200 LeuLeuValProThnGlnpheValaGlAlaIlelleglylsGlnlyalathrlleay 219
QY 721 AACATCATTAAGACACCCAGTCCGGGTAGATTCATATGAAAGAAAGAACTGGAGCT 780
Db 220 AsnIleThrlysglnthrGlnserlysrleaservalhIaArglyselunsnlaGlYAla 239
QY 781 GCAGAGAGGCTGTACCATGCATGCCACCCAGAGGGGACTTCTGAACATGCGGCATG 840
Db 240 AlaGlylsSerIleThrIleleuserthrProglunglythhserAlaIacylsSer 259
QY 841 ATCTTGAATCATGCAGAAAGAGGAGAGTGAACCAAACTAGCCGAGAGAGTTCCTG 900
Db 260 IleleuglntleMehIlelyseluAlaGlnaprllelyspheThrGlnGlnIleProleu 279
QY 901 AAAATCTTGGCACACATAGCGCTTGGTGAAGACTGATTGAAAGAAAGGAGAAATTG 960
Db 280 LysIleleuAlahIsasnasnPheValaGlyArgleuIleGlyLygluGlyArgasnneu 299
QY 961 AAGAAATTTGAACATGAACAGAGGACCAAGATTAACATCTCTTTGCGAGATTGAGC 1020
Db 300 LysylsIleGlnGlnAspThrAspThrlysrIleThrIleSerProleuglntleuThr 319
QY 1021 ATATTAACAACCCGAGAAAGAACCATCATCTGTGAAGGGACAGATTGAGCGCCAGTGT 1080
Db 320 LeuTrysnenProglunArgThrIleThrVallysglyasnValaGlnthCysAlaIysAla 339
QY 1081 GAGATTAAGATTATGAAGAGCTCGTAGGCGCTTTGAAATGATATGATCGCTGTAAAC 1140
Db 340 GlnGlnGlnIleMetlylsylleArgGlnserTyrgluasnnaprllealaserMetasn 359
QY 1141 ACCGACTCC-----GCATCTTC----- 1158
Db 360 LeuGlnAlahIsleuIleProGlyLeuasnleuasnAlaGlnGlyLeuPheProProthr 379
QY 1159 -----TCAGGCTGTAGCCCGCCCATCAACGATTTGGCCGCTTCCCGCAT 1200
Db 380 SerGlyMetProProthrSerGlyProProSerAlaMetThrProProlyrProGln 399
QY 1201 CATCATCTTAATCCAGACAGAGATTTGAATCTCTTCAATCCCAACCCAGCGTGGGC 1260
Db 400 PheGlnGln---SerGlnThrGlnThrValhIeGlnPheIleProAlaleuserValGly 418
QY 1261 GCCATCATGGGAGAAAGGGGGACACATCAATCAACAGCTGGCGAGATTGCCGGAGCTCT 1320
Db 419 AlaIlelleGlylsGlnGlylnhIsIlelysglnleuserArgPheAlaGlyAlaser 438
QY 1321 ATCAAGATTGCCCCCGAGAGGCGCCAGAGCTCAGCGAAAGATGGTCATCATCAACGGG 1380
Db 439 IlelyslleAlaProAlaGlnAlaProAspAlaIyValaArgmetValIleIlethGly 458
QY 1381 CCAACCGAAGCCGATTCAGAGGCCAGAGCGGATCTTTGGGAAACTGAAGAGAAAC 1440
Db 459 ProProGlnAlaGlnPheylsAlaGlnGlyArgIleTyrdlyysIlelysglnGlnasn 478
QY 1441 TTCTTAACCCCAAGAGAGTGAAGCTGGAGCGCATATCAAGATGCCCTCTTCACA 1500
Db 479 PheValaSerProlysglnGlnVallylsleuGlnAlahIsIleArgValProSerPheAla 498
QY 1501 GCTGGCGGGGTATTTGGCAAGAGTGGCAAGACCGTGAAGAACTGAGAACTTAACAGT 1560
Db 499 AlaGlyArgValIleIlelysglysglylysthrValaAsnGlnleuIndlnsnleuserSer 518
QY 1561 GCAAGATCATGTGCTCTGTGACCAACCGCAGATGAAGAAATGAGAGATGATCTCAG 1620
Db 519 AlaGlnValaValaProArgaspGlnThrProAspGlnAsnAspGlnValaVallyls 538
QY 1621 ATTATCGGCACTTCTTGTGCTGACCAAGTGCACAGCGGCAAGATCAGGGAATTTGACAA 1680


```

Db      346  lglmlaGlYAsnProGlyGlyProGlyProGlyGlyArgGlyArgGlyArgGlyGlnGln 366
QY      1426  -----CTGAAGAGGAAACTTCTTAAACCCCAAGA 1457
Db      366  yAsnTrpAsnMetGlyProProGlyGlyLeuGlnGlnPheAsnPheIle----- 382
QY      1458  AGAAGTGAAGCTGAAGAGCCGATATCAGAGTCCCTCTCCACAGTCGCGGGTGAATTGG 1517
Db      383  -----ValProThrGlyGlyThrGlyLeuIleIleGln 393
QY      1518  CAAAGTGGCAAGACCTGACAGAACTGACAGAACTTAACAGTGCAGAAAGTATCGTGGC 1577
Db      393  ylysglyglyglutThrIleLeuSerIleSerGlnGlnSerGlyAlaArgIleGlnGln 413
QY      1578  TCGTACCCAAACGCGCAGATGAATAATGAGGAAGT---ATCGCAAAATTATGGGAC-- 1632
Db      413  nArgAsnProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThrPr 433
QY      1633  -----TTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATTGTACAACA 1682
Db      433  oGlnGlnIleAspTyrAlaArgGlnLeuIleGlnGlnIleGlyGlyProValAsnPr 453
QY      1683  GGTGAAGACGAGCGAGCAAAATACCTTCAGGAGTGGCTTCACAGCGCAGCAAGTGAAG 1742
Db      453  oLeuGlyProProValProHisGlyProHisGlyVal-ProGlyProHisGlyProProG 473
QY      1743  CTCCACAGGACCGCCAGGAGACTCTGCGGAG---CCCTGAAGACCCGAGCGCCGAGAGG 1919
Db      494  -----ProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGly 509
QY      1920  GCGCGGAGAGGTGACGAGGTTTCCAGAA-----CCACGAGCGCCGCC 1964
Db      510  TrpGlyAsnAlaTyrProHisTrpGlnGlnAlaProProAspProAla 526

RESULT 10
US-08-726-160-2
; Sequence 2, Application US/08726160
; Patent No. 5734016
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL POSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,160
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/021,608
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FELLER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 644
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide/Protein
; HYPOTHETICAL: No
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; FEATURE:
; OTHER INFORMATION:
; US-08-726-160-2
; Amino Acid 149 (Xaa) is Met or Ile

Alignment Scores:
Pred. No.: 1.52e-14 Length: 644
Score: 263.50 Matches: 140
Percent Similarity: 37.73% Conservative: 86
Best Local Similarity: 23.37% Mismatches: 237
Query Match: 4.48% Indels: 136
Gaps: 25

US-09-270-437d-8 (1-3283) x US-08-726-160-2 (1-644)
QY      336  CCTCTCTACCTGCGAGTGGAGAGTGTGGATGACTTTTGCTCAATATGGAGACAGTGA 395
Db      10  ProSerSerGlySerAlaGlyGlyGly-----GlyGlyGly 22
QY      396  GAATGTGAACMACTCAACACAGACAGAAACCGCGTTGTCAACGTCACTATGCAAC 455
Db      23  GlyGlyGly-GlyValAsnAspAlaPheLysAspAlaLeuGlnArg-----AlaAr 39
QY      456  AAGAGAAAGAAACAAATAGCATGAGAAAGCTAAGCGGGCATGATTGGAATCTACTC 515
Db      39  GlnIleAlaAlaLysIleGlyLysAspAlaGlyThrSerLeuAsnSerAsnAspTyrGln 59
QY      516  CTTCAGATTCTCTACATCCGAGTGAAGAGTGAAGTCCCTTCGCGCCCTCAGCGAGC 575
Db      59  YTYGlyGlyGlnLysArgProLeuGlnAspGlyAspGlnProAspAlaLysValAl 79
QY      576  CCAAGCGTGGGACCATCTTCCGCGAGCAAGCCAGCCCTGGGGGCACTTCTCAGGC 635
Db      79  aProGlnAsnAspSerPheGlyThrGlnLeu-----ProProMetHisGlnGlnGln 97
QY      636  CAACAGATTGATTTCCCGCTGCGAGATCTGTGTCCTCCACCAAGTTGTTGTCATCAT 695
Db      97  rArgSerVal---MetThrGlnGlnTyrLysValProAspGlyMetValGlyPheIle 116
QY      696  CGAAAGAGAGGCTTGAACCTTAAGAACAATCACTAAGCAAGCAACCAAGTCCCGGTAGAT 755
Db      116  eGlyArgGlyGlyGlnGlnIleSerArgIleGlnGlnGlnSerGlyCysLysIleGln 136
QY      756  CCATGAAAAAGAACTGAGCTGCAGAGAACTGTACACATCCATCCAGCCCGCAGAG 815
Db      136  e---AlaProAspSerGlyGlyLeuProGlnLysSerCys**LeuThrGlyThrProGln 155
QY      816  GGGGACTTCTGAAGCTGCCCATGATTCTTGAATCATGACAGAA----- 861
Db      155  uSerValGlnSerAlaLysArgLeuAspGlnIleValGlnLysGlyArgProAlaPr 175
QY      862  -----GAGGACATGAGAGCCAAACTAGCGGAAGATTCTCTGAAATCTTGGC 911
Db      175  oGlyPheHisHisGlyAspGlyProGlyAsnAlaValGln-----GlnIleMetIle 192

```

[illegible]

Db 494 -----ProGlyProAlaProIseGlyProProAlaProClyAlaProGInGly 509
QY 1920 GGGGGGAAGTCAGCCAGGTTTGCAGAA-----CCACCGAGCCCCGCC 1964
Db 510 TrpGlyAsnAlaTy-PrHisTrpGInGInAlaProProAspProAla 526

RESULT 11
PCT-US94-01782-2
Sequence 2, Application PC/TUS9401782
GENERAL INFORMATION:
APPLICANT: THE GOVERNMENT OF THE UNITED STATES
APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
APPLICANT: HEALTH AND HUMAN SERVICES
TITLE OF INVENTION: NOVEL FUSE BINDING
TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01782
FILING DATE: 22-FEB-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/021,608
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAM S. FEILER
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4063PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 644
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide/Protein
HYPOTHETICAL: No
ORIGINAL SOURCE:
ORGANISM: Human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE: HL60
ORGANELLE:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
OTHER INFORMATION: Amino Acid 149 (Xaa) is Met or Ile
PCT-US94-01782-2

Alignment Scores:
Pred. No.: 1,52e-14 Length: 644
Score: 263.50 Matches: 140
Percent Similarity: 37.73% Conservative: 86

Best Local Similarity: 23.37% Mismatches: 237
 Query Match: 4.48% Indels: 136
 DB: 5 Gaps: 25

US-09-270-437D-8 (1-3283) x PCT-US94-01782-2 (1-644)

```

QY 336 CCTCTCTACCTTCACTAGGAGGTGTGTGATGAGCTTTGGCTCAATATATGGACAGTGA 395
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 10 ProSerSerGlySerIaGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 22
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 396 GAATGTGGAACAGACACACACACACACACACACACACACACACACACACACACAC 455
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 23 GlyGlyGly-GlyValAsnAspAlaPheLysAspAlaLeuGlnArg-----AlaAr 39
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 456 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 515
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 39 gGlnIleAlaAlaLysIleGlyGlyAspAlaGlyThrSerLeuAsnSerAsnAspIle 59
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 516 CTTCAGATTTCTTACATCCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 575
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 59 YTyGlyGlyGlnLysArgProLeuGlnAspGlyLysGlnProAspAlaLysValAl 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 576 CAGAGGTGGAGCACTCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 635
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 79 AProGlnAsnSperPheGlyThrGlnLeu-----ProPheMetHisGlnGlnLase 97
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 636 CAGACAGATTGATTTCCCGCTGCGAGATCCGTGGTCCCAACCCAGTTTGGTGGCCATCAT 695
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 97 rArgSerVal---MetThrGlnGlnIleLysValAlaProAspGlyMetValGlyPheIle 116
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 696 CGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 755
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 116 eGlyArgGlyGlyGlnGlnIleSerArgIleGlnGlnGlnGlnGlnGlnGlnGlnIle 136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 756 CCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 815
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 136 e---AlaProAspSperGlyGlyLeuProGlnArgSerCys**LeuThrGlyThrProG 155
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 816 GGGAGCTTCTGAACATGCCATGATCTTGAATCATGACAGAA----- 861
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 155 UserValGlnSerAlaLysArgLeuLeuAspGlnIleValGlnLysGlyArgProAlaPr 175
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 862 -----GAGCGACATGAGACCAACTAGCCGAGAGAGATTCCTGAAATCTTGGC 911
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 175 oGlyPheHisHisGlyAspGlyProGlyAlaAlaValGln-----GluIleMetI 192
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 912 ACACAAATGGCTTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 971
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 192 eProAlaSerLysAlaGlyLeuValIleGlyLysGlyGlyGlnThrIleLysGlnLeuG 212
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 972 ACATGAAACAGAGCAACAGATTAACATCTGATCTTGGAGATTTGAGCATATACACC 1031
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 212 nGlnArgAlaGlyAlaLysMetValMet-----IleGlnAsp-----GlyPr 226
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1032 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1076
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 226 oGlnAsnThrGlyAlaAspLysProLeuArgIleThrGlyAspProGlyrLysValGlnG 246
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1077 TGCTGAGATAGAGATTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1136
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 246 nAlaLysGlnMetValLeuGlnLeuIleArgAsp----- 257
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1137 TAAACCCACTCCGATCTTCTCCAGCTGTACCCCATACACAGTTTGGCCCGTTCC 1196
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 258 -----GlnGlyGlyPheArgGlnVal-----ArgAsnLysIleGly----- 269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1197 GCATCACTACTTATCCAGAGAGAGAGATTTGAATCTTCTCCACACAGAGCTGT 1256
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 270 -----SerArgIleGlyLysnGlnGlyLysValAlaProIleProArgPheAla 287
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1257 GGGGCGCATCATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1316
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 287 lGlyIleValIleGlyArgAsnGlyLysMetIleLysIleGlnAsnAspAlaGlyVal 307
  
```

```

QY 1317 CTCTATCAAGATTGCCCTTCCGAGAGCCCAAGCTCAAGAGAGAGAGATGTCATCATC 1376
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 lArgIleGlnPheLysProAspAspGly---ThrThrProGlnArgIleAlaGlnIleTh 326
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1377 CGGGCCACCGGAA-----GCCCATTTCAAGGCCAG----- 1407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 326 rGlyProProAspArgCysGlnHisAlaIleGlnIleIleThrAspLeuLysSerVal 346
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1408 ----- 1425
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 346 lGlnAlaGlyAsnProGlyGlyProGlyProGlyArgGlyArgGlyArgGlyGlnG 366
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1426 -----CTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1457
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 366 YAsnTrpAsnMetGlyProProGlyGlyLeuGlnGlnPheAsnPheIle----- 382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1458 AGAATGAAAGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1517
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 383 -----ValProThrGlyLysThrGlyLeuIleIleG 393
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1518 CAAAGGTGCAAGACCTGTAACGAACTGACAGAACTTAACAGTGAAGATCATGTGCC 1577
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 393 YLysGlyGlyGlnThrIleLysSerIleSerGlnGlnSerGlyAlaArgIleGlnLeuG 413
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1578 TCGTACCAAAAGCCAGATGAAATGAGAAAGT---ATCGTCAAGATTAATGCGGAC-- 1632
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 413 nArgAsnProProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThrPr 433
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1633 -----TTCTTGTACCAAGACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1682
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 433 oGlnGlnIleAspIleAlaArgGlnLeuIleGlnGlnGlnGlnGlnGlnGlnGlnIle 453
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1683 GGTGAAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1742
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 453 oLeuGlyProProValProHisGlyProHisGlyVal-ProGlyProHisGlyProProG 473
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1743 CTCCACAGGAGCAGCAGCAAAAGAGATGATGATGATGATGATGATGATGATGATGATG 1802
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 473 lYProProGly-ProGly-----ThrPro-----Met 481
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1803 AGACCAAAAGAGCAGCAGCAGATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1862
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 482 GlyProIleAsnProAlaProIleAsnProGlyPro----- 493
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1863 CTGCGAGAGGCGCAGGAGACTGTGCGAGG---CCCTGAGAAACCCAGGCGCAGAGAG 1919
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 494 -----ProGlyProAlaProHisGlyProProAlaProIleAlaProGlnGly 509
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1920 GGGGAGAGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1964
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 510 TrpGlyAsnAlaIleTyProHisIleTrpGlnGlnAlaIleProProAspProAla 526
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  
```

RESULT 12
 US-08-021-608D-10
 ? Sequence 10, Application US/08021608D
 ? Patent No. 5580760
 ? GENERAL INFORMATION:
 ? APPLICANT: LEVENS, DAVID L., DUNCAN,
 ? APPLICANT: ROBERT C., AND AVIGAN, MARK I.
 ? TITLE OF INVENTION: NOVEL FUSE BINDING
 ? NUMBER OF SEQUENCES: 24
 ? CORRESPONDENCE ADDRESSES:
 ? ADDRESSEE: MORGAN & FINNEGAN
 ? STREET: 345 PARK AVENUE
 ? CITY: NEW YORK
 ? STATE: NEW YORK
 ? COUNTRY: USA
 ? ZIP: 10154
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: FLOPPY DISK
 ? COMPUTER: IBM PC COMPATIBLE

QY	1743	CTCCACAGGACCCAGCAAAACACGATGAATGATGCCCTTCCAAACCTGACAGATG	1802
Db	472	LYPProGly-ProGly-----	ThiPro-----Met 480
QY	1803	AGACCAAAACGACGACCCAGATCGGGAGCAAAACCAAGCCATCTGAGATGAGAAGT	1862
Db	481	GLYProTYrAsnProAlaProTYrAsnProGlyPro-----	-----492
QY	1863	CTGGGAGAGCGCGCCAGGACCTCTGCCGAGG-----CCCTGAGAAACCCAGGGCCGAGAGG	1919
Db	493	-----ProGlyProAlaProHisGlyProProAlaProTYrAlaProGlnGly	508
QY	1920	GGCGGGAGAGGTCAAGCCAGGCTTTCCAGAA-----CCACGAGACCCCGGCC	1964
Db	509	TrpIlyAsnAlaTYrProHisTrpGlnGlnGlnAlaProProAlaProProAla	525

```

Db      286  lGlytleVallelglYArGAnGlylMelelleYslYsrlleGlnsnaSPalaGlyVa 306
QY      1317  CTTATCAAGATTGCCCCCTGGGAAGGCCAGCTGACGAAAGATGTCATCATCAC 1376
Db      306  lArGlelnePhesProAspArgly---ThrThrProGlnArglleAlaGlnleThr 325
QY      1377  CGGGCCACCGGAA--GCCCGATTCAAGCCGAG----- 1407
Db      325  rGlyProProAspArgCysGlnHlsAlaAlaGlnlelleThrAspleuleuAArgSerVa 345
QY      1408  -----GGACGATCTTTGGGAAA----- 1425
Db      345  lGlnAlaGlyAsnProGlyGlyProGlyProGlyArgGlyArgGlyArgGlyGlnG 365
QY      1426  -----CTGAAGAGGAAACTCTTTAACCCCAAGA 1457
Db      365  YAsnTrpAsnMetGlyProProGlyGlyLeuGlnGlnlupheAsnPhelle----- 381
QY      1458  AGAAGTGAAGCTGGAAGCGCATATGAGATGCGCTCTTCCACAGCTGGCGGGGATTGG 1517
Db      382  -----ValProThrGlyLysThrGlyLeuIlelleG 392
QY      1518  CAAAGTGGCAAGACCGTGAAAGAACTGCAAGACTTAAACAGTGCAGAAAGTATGTCGC 1577
Db      392  YLysGlyGlyGlnThrIleYSerlleSerGlnGlnInseryAlaArglleGlnleuG 412
QY      1578  TCGTGAACCAAGCCGATGAATAATGAGAAAGT--ATGTCAGAAATTATGGGGAC-- 1632
Db      412  nArgAsnProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThrPr 432
QY      1633  -----TTCTTGTACGACCTGCAAGCCGACAGCCAGATCAAGGGAATTTACAA 1682
Db      432  oGlnGlnIleAspTYAlaArgGlnleuIleGlnGlnYslleGlyGlyProValAsnPr 452
QY      1683  GGTGAAGCAGCAGAGAGCAAGAAATACCCCTGAGGAGTGGCTTCAAGCGCAGCAAGTAG 1742
Db      452  oLeuGlyProProValProHisGlyProHisGlyVal-ProGlyProHisGlyProProG 472
QY      1743  CTCGCCAGGACCCAGCAAAACAGAGATGATGACCTTCCACACCTGACAGAAATG 1802
Db      472  lYProProGly-ProGly-----ThrPro-----Met 480
QY      1803  AGACCAAGCCAGCCAGCCAGATCGGGAGCAACCAAGACATGTAGAGATGAGAAAT 1862
Db      481  GLyProTYAsnProAlaProTYAsnProGlyPro----- 492
QY      1863  CTCGCGAGCGCGCCAGGACTTGCAGAG--CCCTGAGAACCCGAGGGCCGAGAGG 1919
Db      493  -----ProGlyProAlaProHisGlyProProAlaProTYAlaProGlnGly 508
QY      1920  GGGCGGAGAGTCAAGCCAGGTTGCCGAA-----CCACCGAGCCCGCGC 1964
Db      509  TrpGlyAsnAlaTYProHisTrpGlnGlnAlaProProAspProAla 525

```

RESULT 14
PCT-US94-01782-10
Sequence 10, Application PC/TUS9401782

GENERAL INFORMATION:
APPLICANT: THE GOVERNMENT OF THE UNITED STATES
APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
APPLICANT: HEALTH AND HUMAN SERVICES
TITLE OF INVENTION: NOVEL FUSE BINDING
TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK

```

COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01782
FILING DATE: 22-FEB-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/021,608
FILING DATE: 22-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAM S. FEILER
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4063PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 643
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide/Protein
HYPOTHETICAL: No
ORIGINAL SOURCE: Human
ORGANISM: Human
STRAIN:
INDIVIDUAL ISOLATE:
DEVELOPMENTAL STAGE:
HAPLOTYPE:
TISSUE TYPE:
CELL TYPE:
CELL LINE: HL60
ORGANELLE:
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PCT-US94-01782-10
Alignment Scores:
Pred. No.: 2.06e-14 Length: 643
Score: 262.00 Matches: 140
Percent Similarity: 37.56% Conservative: 85
Best Local Similarity: 23.37% Mismatches: 237
Query Match: 4.45% Indels: 137
Gaps: 25
US-09-270-437D-8 (1-3283) x PCT-US94-01782-10 (1-643)
QY      336  CCTCCCTCACTGAGAGGAGTGTGATGACTTTTGCTCAATATGAGACATGGA 395
Db      10  ProSerSerIysAlaGlyGlyGly-----GlyGlyGly 22
QY      396  GAATGTGAAACAAGTCAACAACAACAAGAACCGCGCTGTCAAGTCAATATGCAAC 455
Db      23  GlyGlyGlyGlyGlyValAsnAsnSPalaPhelyAspAlaLeuGlnArg-----AlaAr 39
QY      456  AAGAGAAAGAAATAATGCGATGAGAGAAAGCTAAGGGGATGAGTTGAGAACTACTC 515
Db      39  glnlleAlaAlaYslleGlyGlyAspAlaGlyThrSerLeuAsnSerAsnAspTYrG 59
QY      516  CTTGAAGATTTCCACATCCCGATGAAGAGTCCCTTCGCCCCCTCAGCGAGC 575
Db      59  YTYrGlyGlyGlnYsarGProLeuGlnAspGlyAspGlnProAspAlaYslYValAl 79
QY      576  CCAGCGTGGGAGCACTCTTCCCGGAGCAAGCCGCCCTGGGGGCACTTCTCAGGC 635
Db      79  aProGlnAsnAspSerPheGlyThrGln-----LeuProProwMetHisGlnGlnG 96

```



```

QY 636 CAGACGATGATTTCCCGCTGGAGTCTGTCCTGCTCCACCAGTTTGTGTGCTCATCAT 695
Db |||
QY 96 nArgSerVal---MetThrGluGluIuTrIySerValProAspGlyMetValGlyPheIle11 115
QY 696 CGAAGAGAGGGGCTTGACCATTAAGAACTCTACTAAGCAGACCCACTCCGGGTAGANT 755
Db |||
QY 115 eGlyIaArgGlyGlyGluGlnIleSerArgIleGlnGlnIuSerGlyCyAlaIleGln11 135
QY 756 CCATGAGAAAGAGAACTCTGAGCTGCAGAGAGAGCTGTCACTCATTCACACCCGCA 815
Db |||
QY 135 e---AlaProAspSerGlyGlyLeuProGluArgSerCys---LeuThrGlyThrProG1 154
QY 816 GGGGACTTCTGAGACATGCCCATATCTTGAATTCATGCAGAAA----- 861
Db |||
QY 154 uSerValGlnSerAlaIySarGlyLeuAspGlnIleValGluIySgIyArgProAlaPr 174
QY 862 -----GAGCGAGATGAGCAACCAACTAGCCCAAGAGATCTCTGAAATTTGGC 911
Db |||
QY 174 oGlyPheHisIseIyAspGlyProGlyAsnAlaValGln-----GluIleMet11 191
QY 912 ACACAAATGGCTTGGTGAAGACTGATGAGAAAGAGAGAAATTTGAAGAAATGA 971
Db |||
QY 191 eProAlaSerIySaIaGlyLeuValIleGlyIySgIyGluThrIleIySgIleuG1 211
QY 972 ACATGAAACAGGAGCAACGATTAACATCTCATCTTGGAGGATTTGACATATACACCC 1031
Db |||
QY 211 nGluIaArgAlaGlyAlaIySmetValMet-----IleGlnAsp-----GlyPr 225
QY 1032 GGAAGAGAAC-----ATCACTGTGAAGGGCAGCTGAGCCCTGTGCAG 1076
Db |||
QY 225 oGlnAsnThrGlyAlaAspIySProLeuArgIleThrGlyAspProIyIySValGlnG1 245
QY 1077 TGCTGAGATAGAGATTATGAGAGAGCTGCTGAGAGCTTGTGAATGATATGCTGGCTGT 1136
Db |||
QY 245 nAlaIyGluMetValIleuGluIeIyAsp----- 256
QY 1137 TAAACCCACTCCGGATACTTCTTCACGCTGACCCCATCACCAAGTTGGCCGTTCCC 1196
Db |||
QY 257 -----GlnGlyIyPheArgGluVal-----ArgAsnIuTrIyGly----- 268
QY 1197 GCATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCACTCCCAACCGAGCTGT 1256
Db |||
QY 269 -----SerArgIleGlyGlyAsnGluGlyIleAspValProIleProArgPheAlaVa 286
QY 1257 GGGCGCCATCATCGGAGAGAGGAGACACATCAACACACTGGCCGAGATTCGCCGAGC 1316
Db |||
QY 286 IGIyIleValIleIleGlyIyArgSngIyGluMetIleIySylIeGlnAsnAspIaGlyVa 306
QY 1317 CTCTATCAGATGTTCCCTCGGAGGCCCAAGCTCAGCGAAAGATGTCTATCATCAG 1376
Db |||
QY 306 IArgIleGlnPheIySProAspSgIy---ThrThrProGluArgIleAlaGlnIleTh 325
QY 1377 CGGGCCACCGGAA---GCCAGTTCAAGGCCAG----- 1407
Db |||
QY 325 rGlyProProAspArgCysGlnIleSaIaAlaGluIleIleThrAspLeuIuArgSerVa 345
QY 1408 -----GAGCGGATCTTGGGAAA----- 1425
Db |||
QY 345 IGIaIaGlyAsnProGlyGlyProGlyIyArgGlyIyArgGlyIyArgGlyIyGlnG1 365
QY 1426 -----CTGAAGAGAGAAACTCTTTAAACCCCAAGA 1457
Db |||
QY 365 yAsnTrpAsnMetGlyProProGlyIyGlyLeuGlnIuPheAsnPheIle----- 381
QY 1458 AAGAGTGAAGCTGGAAGCGCATATCAGATGCCCTCTTCCACAGCTGCGCGGTATGG 1517
Db |||
QY 382 -----ValProThrGlyIySthGlyLeuIleIleG1 392
QY 1518 CAAAGGTGCAAGACCGTGAAGCAAGAACTTAACCAAGTGAAGATCATCTGTC 1577
Db |||
QY 392 IySgIyGlyIuThrIleIySserIleBerGlnGlnSerGlyAlaArgIleIleIuG1 412
QY 1578 TGTGATCAACCAACCGATGAAATGAGGAAGTG---ATCGTCAAGATTATCGGGCAC-- 1632

```

```

Db 412 nArgAsnProProProAsnAlaAspProAsnMetIySLeuPheThrIleArgGlyThrPr 432
QY 1633 -----TTCCTGTAGCAGACTGCAGAGCGGCAAGATCAGGAAATTTGCANCA 1682
Db |||
QY 432 oGlnGlnIleAspIyRAlaArgGlnIleuIeGluGluIySylIeGlyIyProValAsnPr 452
QY 1683 GGTGAAGACGACGAGCAGAAATACCTCAAGGAGTGGCTTCACAGCGCAGCAAGTGAGG 1742
Db |||
QY 452 oLeuIyProProValProHisGlyProHisGlyValProIyProHisGlyProProG 472
QY 1743 CTCCACAGGACCGAGCAACCAACAGGATGATGAGCCCTTCCAAACCTTCAAGATG 1802
Db |||
QY 472 IyProProGlyProGly-----ThrPro-----Met 480
QY 1803 AGACCAAGCAGCAGCCAGATCGGAGCAACCAACCAAGCATCTGAGGAATGAGACT 1862
Db |||
QY 481 GlyProIyAsnProAlaProIyIyAsnProGlyPro----- 492
QY 1863 CTGGGAGGCGGCGCAGAGACTCTGCCAGG---CCCTGAGAAACCCAGGCGCAGAGAG 1919
Db |||
QY 493 -----ProGlyProAlaProHisGlyProProAlaProIyAlaProGlnGly 508
QY 1920 GCGGGAAGGTACGACGATTTGCCAGAA-----CCACCGAGCCCCGCC 1964
Db |||
QY 509 TrpIyAsnAlaIyTrProHisTrpGlnGlnIaIaProProAspProAla 525

RESULT 15
US-08-021-608D-8
; Sequence 8, Application US/08021608D
; Patent No. 5580760
;
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,608D
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-6849
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide/Protein
; HYPOTHETICAL: No
; ORIGINAL SOURCE:

```

ORGANISM: Human
CELL LINE: HL60
US-08-021-608D-8

Alignment Scores:

	8.21e-14	Length:	590
Pred. No.:	255.00	Matches:	116
Score:	38.52%	Conservative:	72
Percent Similarity:	23.77%	Mismatches:	178
Best Local Similarity:	4.33%	Indels:	122
Query Match:	1	Gaps:	21

US-09-270-437D-8 (1-3283) x US-08-021-608D-8 (1-590)

```

QY GTCCCAACCCAGTTTGTGTCATCCGAAAGAGGCTTGACATTAAGACATC 726
   |||||
Db ValProAspGlyMetValGlyPheIleIleGlyArgGlyGlnGlnIleSerArgIle 110
QY 727 ACTAAGCAGACCCAGTCCCGGTAGATTCATAGAAAAGAACTCTGGAGCTGCAG 786
   |||||
Db 111 GlnGlnIleSerGlyCysIleIleGlnIle--AlaProAspSerGlyGlyLeuProGlu 129
   |||||
QY 787 AAGCTGTCACCATCCATCCACCCAGAGGGGAACTTGAAGCATGCGCATGATCTT 846
   |||||
Db 130 ArgSerCysMetLeuThrGlyThrProGluSerValGlnSerAlaIleArgLeuLeuAsp 149
   |||||
QY 847 GAATCATGACAGAAA-----GAGCAGATGAGACCAACTA 882
   |||||
Db 150 GlnIleValGlnIleGlyArgProAlaProGlyPheHisIleGlyAspGlyProGlyAsn 169
   |||||
QY 883 GCGGAGAGAGATCTCTGAAATCTTGGACACACATGCTTGTGGAAAGATGATGGA 942
   |||||
Db 170 AlaValGlnIle-----GlnIleMetIleProAlaSerIleValGlyLeuValIleGly 186
   |||||
QY 943 AAGAAGAGGAGAAATTTGAAGAAATTGAACATGAACAGGACCAAGATTAACATCTCA 1002
   |||||
Db 187 LysGlyGlyGlnThrIleLysGlnLeuGlnGlnIleValGlyValLysMetValMet--- 205
   |||||
QY 1003 TCTTTGACAGATTGTGACATATACAAACCGGAAAGACC-----ATCACT 1047
   |||||
Db 206 ---IleGlnAsp-----GlyProGlnAsnThrGlyAlaAspLysProLeuArg 220
   |||||
QY 1048 GTGAAGCGGACAGTTGAGGCTGTGCGCATGCTGCAATGAGATTAAGAAAGCGCT 1107
   |||||
Db 221 IleThrGlyAspProGlyLysValGlnGlnIleAlaValGlnIleValGlnIleValGln 240
   |||||
QY 1108 GAGGCTTTGAATAATGATATGCTGGCTTAAACCACTCCGATATCTTCCAGCTG 1167
   |||||
Db 241 Asp-----GlnGlyGlyPheArgGlnVal 248
   |||||
QY 1168 TACCCCATCACAGTTTGGCCCGTCCGATCATCATCTTATCCAGACGAGATT 1227
   |||||
Db 249 ---ArgAsnGlnIleGly-----SerArgIleGlyGlnIleGlnIleGlnIle 261
   |||||
QY 1228 GTGAATCTTTCATCCCAACCGAGCTGTGGCGCCCATCATCGGAGAAAGAGGCGCAC 1287
   |||||
Db 262 IleAspValProIleProArgPheAlaValGlyIleValIleGlyArgAsnGlyGlnMet 281
   |||||
QY 1288 ATCAAAAGCTGGGAGATTCGCGGAGCTCTATCAAGATTGCCCTGCGGAAAGGCCCA 1347
   |||||
Db 282 IleLysLysIleGlnAsnAspIleGlyValArgIleGlnPheLysProAspAspGly--- 300
   |||||
QY 1348 GAGCTCAGCAGAAAGATGTCATCATCAACCGGCGCACCGGAA---GCCAGTTCAAGGCC 1404
   |||||
Db 301 ThrThrProGluArgIleAlaGlnIleThrGlyProProAspArgCysGlnHisAlaIle 320
   |||||
QY 1405 CAG----- 1407
   |||||
Db 321 GlnIleIleThrAspLeuLeuArgSerValGlnIleAsnProGlyGlyProGlyPro 340
   |||||
QY 1408 ---GAGCGGATCTTGGGAAA-----CTG 1428
   |||||
Db 341 GlyGlyArgGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProProGlyGlyLeu 360

```

```

QY 1429 AAGAAGAAACTCTTTAATCCCAAGAAAGAAAGCTGAAAGCCATATCAGATG 1488
   |||||
Db 361 GlnGlnPheAsnIle-----Val 367
   |||||
QY 1489 CCTCTTCCACAGCTGCGCGGTGATTTGCAAAAGTGGCAAGACCGTAACGATCGAG 1548
   |||||
Db 368 ProThrGlyLysThrGlyLeuIleIleGlyLysGlyGlnIleLysSerIleSer 387
   |||||
QY 1549 AACTTAACGATGCGAGAAAGTCATGCTGCTGTCAGCAAAACGCGATGAAATAGAGAA 1608
   |||||
Db 388 GlnGlnSerGlyAlaArgIleGlnLeuGlnArgAsnProProProAsnAlaAspProAsn 407
   |||||
QY 1609 GTG---ATCGTCAGAAATATCGGCAC-----TTCTTGTCTAGCAAGACTGCA 1653
   |||||
Db 408 MetLysLeuPheMetIleArgGlyThrProGlnGlnIleAspIleArgGlnIle 427
   |||||
QY 1654 CAGCGCAAGATCAGAGGAAATTTGACAAAGGTGAAGCAGAGCAGAGCAAGAAATACCTCAG 1713
   |||||
Db 428 GlnGlnLysIleGlyGlyProValAsnProLeuGlyProProValProHisGlyProHis 447
   |||||
QY 1714 GAGTCGCTTCACAGCGCAGCAAGTGAAGCTCCACAGGACCAAGCAAAACAGCATGA 1773
   |||||
Db 448 GlyVal-ProGlyProHisGlyProProGlyProProGly-ProGly----- 462
   |||||
QY 1774 ATGTAGCCCTTCCACACCTGACAGATGAGACCAAGCCAGCCAGATCGGAGCA 1833
   |||||
Db 463 ---ThrPro-----MetGlyProTyAsnProAlaProTyAsnProG 476
   |||||
QY 1834 AACCAAAACATCTGAGAGATGAGAAAGTTCGCGGAGGCGGCAAGTCTGCGGAGG 1892
   |||||
Db 476 LysPro-----ProGlyProAlaProHis 484
   |||||
QY 1893 ---CCCTGAAGACCCGAGGCGGAGGAGGCGGAGAGGTGAGCGATTTGCCAGAA- 1949
   |||||
Db 484 LysProProAlaProTyArgIleProGlnGlyTrpGlyAsnAlaTyProHisTrpGlnGln 504
   |||||
QY 1950 ---CCACCGAGCCCGCC 1964
   |||||
Db 504 ImlAlaProProAspProAla 510

```

Search completed: July 23, 2004, 11:20:10
Job time : 65.5087 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 11:15:41 ; Search time 156.647 Seconds

(Without alignments)
13126.119 Million cell updates/sec

Title: US-09-270-437d-8

Perfect score: 5886

Sequence: 1 ggcagcgagagagcgcagga.....aaccttgaaatgttattc 3283

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters: 2576884

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/cgn2_1/USPRO.spool/p/US09270437/runat_23072004.095742.27560/app.query.fasta_1.5562
-DB=Published Applications AA -QPM=fastan -SUFFIX=rabp -MINMATCH=0.1
-LOOPC=0 -LOOEXT=0 -UNITS=dics -START=1 -END=-1 -MATRIX=blonsum62
-TRANS=human40.cdi -LIST=45 -DICALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSEBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	2838	48.2	556	14	US-10-097-340-147	Sequence 147, App
2	2838	48.2	556	16	US-10-646-593-182	Sequence 182, App
3	2808	47.7	620	9	US-09-764-864-1116	Sequence 1116, App
4	2745.5	46.6	587	15	US-10-313-986-501	Sequence 501, App
5	2584.5	43.9	555	15	US-10-262-445-40	Sequence 40, App1
6	1940	33.0	577	15	US-10-313-986-500	Sequence 500, App
7	1934	32.9	577	9	US-09-873-637-2	Sequence 2, App11
8	1886	32.0	589	15	US-10-313-986-486	Sequence 486, App
9	1875	31.9	579	9	US-09-735-705-348	Sequence 348, App
10	1875	31.9	579	9	US-09-850-716A-348	Sequence 348, App
11	1875	31.9	579	9	US-09-897-778-348	Sequence 348, App
12	1875	31.9	579	9	US-09-897-778-446	Sequence 446, App
13	1875	31.9	579	9	US-09-897-778-449	Sequence 449, App
14	1875	31.9	579	12	US-10-007-700-448	Sequence 448, App
15	1875	31.9	579	12	US-10-007-700-446	Sequence 446, App
16	1875	31.9	579	12	US-10-007-700-449	Sequence 449, App
17	1875	31.9	579	14	US-10-117-982-348	Sequence 348, App
18	1875	31.9	579	14	US-10-117-982-446	Sequence 446, App
19	1875	31.9	579	14	US-10-117-982-449	Sequence 449, App
20	1875	31.9	579	14	US-10-117-982-480	Sequence 480, App
21	1875	31.9	579	15	US-10-313-986-348	Sequence 348, App
22	1875	31.9	579	15	US-10-313-986-446	Sequence 446, App
23	1875	31.9	579	15	US-10-313-986-449	Sequence 449, App
24	1875	31.9	579	15	US-10-313-986-480	Sequence 480, App
25	1870	31.8	586	9	US-09-850-716A-427	Sequence 427, App
26	1870	31.8	586	12	US-10-007-700-427	Sequence 427, App
27	1870	31.8	586	12	US-09-897-778-427	Sequence 427, App
28	1870	31.8	586	14	US-10-117-982-427	Sequence 427, App
29	1870	31.8	586	15	US-10-313-986-427	Sequence 427, App
30	1868	31.7	579	9	US-09-735-705-176	Sequence 176, App
31	1868	31.7	579	9	US-09-850-716A-176	Sequence 176, App
32	1868	31.7	579	9	US-09-897-778-176	Sequence 176, App
33	1868	31.7	579	12	US-09-466-396A-176	Sequence 176, App
34	1868	31.7	579	12	US-10-007-700-176	Sequence 176, App
35	1868	31.7	579	14	US-10-117-982-176	Sequence 176, App
36	1868	31.7	579	15	US-10-313-986-176	Sequence 176, App
37	1865	31.7	579	14	US-10-117-982-484	Sequence 484, App
38	1865	31.7	579	15	US-10-313-986-484	Sequence 484, App
39	1326	22.5	422	16	US-10-408-765A-2088	Sequence 2088, App
40	812.5	13.8	261	9	US-09-764-864-1114	Sequence 1114, App
41	803	13.6	171	9	US-09-764-864-1119	Sequence 1119, App
42	782.5	12.3	250	9	US-09-764-864-1532	Sequence 1532, App
43	750	12.7	171	9	US-09-764-864-1536	Sequence 1536, App
44	676.5	11.5	192	9	US-09-764-864-1117	Sequence 1117, App
45	357	6.1	81	14	US-10-117-982-476	Sequence 476, App

ALIGNMENTS

RESULT 1
US-10-097-340-147
Sequence 147, Application US/10097340
Publication No. US20030087250A1
GENERAL INFORMATION:
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAVAPU
APPLICANT: Sebastian HOESCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G. KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification, Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030

```

; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 147
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-097-340-147

Alignment Scores:
Pred. No.: 4,29e-221 Length: 556
Score: 2838.00 Matches: 556
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.22% Indels: 0
Db: 14 Gaps: 0

US-09-270-437d-8 (1-3283) x US-10-097-340-147 (1-556)
QY 70 ATGATGAACAAGCTTACATCGGGAACCTGAGCCCGCGATCCGCGACGACCTCGG 129
Db 1 MetMetAsnlyseuIyrIleGlyAsnleuSerProAlaValThrAlaAspAsnleuArg 20
QY 130 CAGCTCTTTGGGACAGAGCTGCCCCCTGGCGGACAGGTCTCTGTTGAAGTCCGGCTAC 189
Db 21 GlnLeuPheGlyAspArglyLeuProLeuAlaGlyGlnValLeuLeuIysSerGlyTyr 40
QY 190 GCCTTGTGAGTACACCCCGACGAACTGGGGCATCGCGCCATCGAGACCTCTCGGAT 249
Db 41 AlaPheValAspIyrProAspGlnAsnTTPAlaIleArgAlaIleGlnThrLeuSerGly 60
QY 250 AAGTGAATTGCATGGGAAATCATGGAAGTATTACTCACTCTCTTAAAAAGCTAAG 309
Db 61 LysValGlnLeuHisGlyIysIleMetGlnValAspIyrSerValSerIysIysLeuArg 80
QY 310 AGCAGGAAATTCAGATTGGAACATTCCTCTCTCACTGCAAGTGGAGGTGGATTGGA 369
Db 81 SerArgIysIleGlnIleArgAsnIleProIroHisIleuGlnIyrGlnValLeuAspGly 100
QY 370 CTTTGGCTCAATATGGGACAGTGGAGATGTGGACCAAGTCAACACAGACAGAAACC 429
Db 101 LeuLeuAlaGlnIyrGlyIhrValGlnAsnValGlnIleValAsnThrAspIhrGlnThr 120
QY 430 GCGGTGTCAACGTCATATGCAACAAAGAGAAAGAAATATAGCCATGAGAACTTA 489
Db 121 AlaValAlaAsnValThrIyrAlaThrArgGlnGlnIleAlaIleAlaMetGlnIysLeu 140
QY 490 AGCGGGCATCAGTTGGAACCTACTCTTCAAGATTTCCTACATCCCGGATGAAAGAGT 549
Db 141 SerGlyIhrIleGlnPheGlnAsnIyrSerPheIysIleSerIyrIleProAspGlnIleVal 160
QY 550 AGCTCCCTTCGCCCCCTCAAGCGAGCCAGCGTGGGAGCACTCTTCCGGAGCAAGGC 609
Db 161 SerSerProSerProProGlnIhrGlnAlaGlnIhrGlnIysPheIysSerSerIhrGlnIleGly 180
QY 610 CAGCCCTCTGGGGACCTTCTCAGGCGCAACAGATTGATTTCCGCTGCGGATCTGGTC 669
Db 181 HisAlaProGlyIhrIhrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
```

```

QY 670 CCAACCCAGTTTGTGGTGCATCATCGGAAGAAGAGGCTTGACCATTAAGAATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleIleGlyIysGlnIysLeuThrIleIysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGCTGATATTCATATGAAAAAGAACTCTGGAGCTCGACAGAG 789
Db 221 LysGlnThrGlnIhrSerArgValAspIleHisArgIysGlnAsnSerGlyAlaAlaGlnIys 240
QY 790 CCGTACCATTCATCGACCCCGAGAGGGACCTTGAAGCATGCGCATGATTTGAA 849
Db 241 ProValThrIleHisAlaIhrProGlnIyhrSerGlnIaCyArgMetIleIeuGln 260
QY 850 ATCATGCAAAAAGGCGATGAGCAAACTAGCCGAAAGAGATTCTCTGAAAATCTTG 909
Db 261 IleMetGlnIysGlnAlaAspGlnIhrIysIleuAlaGlnIleIleProLeuIysIleLeu 280
QY 910 GCACCAATGGCTGTGGTGAAGATGATTTGAAAAAGAGGAGAAATTTGAAAAATT 969
Db 281 AlaHisAsnGlyLeuValGlyArgIleuIleGlyIysGlnIyhrAsnIleuIysIle 300
QY 970 GAACATGAAAACAGGGACCAAGATAACAATCTCATCTTTGACAGATTGACATATACAAC 1029
Db 301 GlnHisGlnIhrGlyThrIysIleThrIleSerSerIeuGlnAspIeuSerIleIyrAsn 320
QY 1030 CCGAAGAAACCATCATCTGTGAAGGCAACAGTTGAGGCTTGCCAGTCTGAGATGAG 1089
Db 321 ProGlnArgIhrIleThrValIysGlyIhrValGlnIaCyAlaSerAlaGlnIleGln 340
QY 1090 ATTATGAAAGAACTCGCGTGAAGGCTTTGAAATGATCTGGCTTTAAACCCACTCC 1149
Db 341 IleMetIysIysLeuArgGlnAlaPheGlnAsnAspMetLeuAlaValAsnThrHisSer 360
QY 1150 GGATCTCTCCAGCCTGACCCCATCACAGTTTGCCCGCTCCCGCATCATCTCT 1209
Db 361 GlyIyrPheSerSerIeuIyrProHisIleGlnPheGlyProPheProHisIleHisSer 380
QY 1210 TATCCAGACAGAGATTGTGAATCTTTCATCCCAACCGAGCTGTGGCGCATCATAC 1269
Db 381 TyrProGlnGlnIleValIleValAsnLeuPheIleProThrGlnAlaValGlyAlaIle 400
QY 1270 GGGAGAAAGGGGCGACATCAAAACAGCTGGGAGATTGGCGGACCTCTTCAAGATT 1329
Db 401 GlyIysIysGlyAlaHisIleIysGlnIleuAlaArgPheAlaGlyAlaSerIleIysIle 420
QY 1330 GCGCCCTGCGAAGGCCCGACGTCAGCGAAAGAGATGTCATCATCAACCGGCGCAACGGAA 1389
Db 421 AlaProAlaGlnIyProAspValSerGlnIhrMetValIleIleThrGlyProProGln 440
QY 1390 GCCCAGTTCAAGGCCCGAGGACGGAATCTTTGGAAACTGAAAGAAGAAACTCTTTAAC 1449
Db 441 AlaGlnPheIysAlaGlnIyArgIlePheGlyIysIleuIysGlnIleuAsnIlePheAsn 460
QY 1450 CCAAAAGAAAGATGAAGCTGGAAGGCGCATATCAGAGTGCCCTTTCACAGCTGCGCG 1509
Db 461 ProIysGlnIleValIlyIysLeuGlnAlaHisIleArgValIProSerSerThrAlaGlyArg 480
QY 1510 GTGATTGGCAAGAGTGGCAAGACCGTGAACGAACGCAAGATTAAACAGTGCAGAAATC 1569
Db 481 ValIleGlyIysGlyIysIhrValAsnGlnIleuGlnIleuIhrSerAlaGlnIleVal 500
QY 1570 ATCGTGCCTGTGATCAACACGCGCATGAAATGAGAAATGATTCGTCAGAAATTAATGG 1629
Db 501 IleValIProArgAspGlnIhrProAspGlnAsnGlnIleValIleValAlaGlyIleIleGly 520
QY 1630 CACTCTTTCTAGCCGACGATGCAACAGCGCAAGATTCAGGGAATTTGACAAACAGTGAAG 1689
Db 521 HisPhePheAlaSerGlnIhrAlaGlnIhrGlyIleArgGlnIleValGlnIleValIys 540
QY 1690 CAGCAGAGCAAGAAATACCTCAGAGGAGTGCCTCACAAGCGCACCAAG 1737
Db 541 GlnGlnIleGlnIysIyrProGlnIyValAlaSerGlnIhrArgSerIys 556
```

RESULT 2

US-10-648-593-182
; Sequence 182, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; PRIOR FILING DATE: 2003-08-26
; PRIOR FILING NUMBER: 60/406,385
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 182
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-182

Alignment Scores:

Pred. No.:	4,29e-221	Length:	556
Score:	2838.00	Matches:	556
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	48.22%	Indels:	0
DB:	16	Gaps:	0

US-09-270-437d-8 (1-3283) x US-10-648-593-182 (1-556)

```
QY 70 ATGATGAACAAGCTTACATGAGACCTGAGCCCGCCGCTACGCGGAGACGACCTCCG 129
Db 1 Methectsnhslyseutyrlleglyashneuserproalvalthralspdsleuarg 20
QY 130 CAGCTCTTTGGGAGCAGAGAGCTGCCCCGCGGAGACAGGTCCTGTAAGTCCGGCTAC 189
Db 21 GlnleuheglsparglysleuProleuhalgilylnalleuLeuysSerGlyTyr 40
QY 190 GCCTTGCTGACATACCCGACACAGAACTGGGCCATCCGGCCATGAGACCTCTCGGGT 249
Db 41 AlaphelvalspyrProaspGlnasntrpAlallearglallegtutlrLeuSerGly 60
QY 250 AAAGTGAATTCATGAGGAAATCATGAGATTGATTACTCAGCTCTTAAAGCTAAG 309
Db 61 LysValGlnleuhsisglysillemeGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGAAATTCAGATTGAAAATCCCTCCTCAGCTGAGTGGAGGTGGTGGATGA 369
Db 81 SerArglyseileglnilleargasnilleProProHisleuGlntrpGluValLeuaspGly 100
QY 370 CTTTGGCTCAATATGGGAGACAGTGGAGAAATGGAACATGCAACACAGACAGAAAC 429
Db 101 LeuLeuhalaglnlyrGlythrValGlnasnValGlnGlnValAsnThrAspThrGlnThr 120
QY 430 GCCGTGTCAACCTCATATGCAACAGAGAGAGAAATAGCCATGAGAGAGCTA 489
Db 121 AlalaValalasnValthrTyrAlaThrAspGlnGlnAlaLysillealnetGluLysLeu 140
QY 490 AGCGGACATCAGTTTGAAGACTAATCTCTTCAGATTCTTACATCCCGAGTAAGAGGT 549
Db 141 SerGlyHisGlnheglsuasnTyrSerPheLysilleSerTyrilleProaspGlnGlnVal 160
QY 550 AGCTCCCTTCGCCCCCTCAGGAGAGCCAGCTGGGGAGCACTCTTCCGGAGAGAGG 609
Db 161 SerSerProserProGlnArgAlaGlnArgLysAspHisSerSerArgGlnGlnGly 180
QY 610 CAGCCCTCGGGGAGCTTCTCAGGCGCAGACAGATTGTTCCCGTGGAGATCCGTGTC 669
Db 181 HisAlaProGlyGlyHisSerGlnAlaGlnGlnLysAspPheProleuArgLysLeuVal 200
QY 670 CCCACCCAGTTTGTGTGTCATCATGAGAAAGAGGGGCTTGACATAAAGACATCACT 729
```

```
Db 201 ProthrGlnPheValGlyAlaIlelleGlyLysGlnGlyLeuThrilleLysasnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGGTAGATATCCATGAGAAAGAACTCGAGCTCAGAGAG 789
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGlnLysSerGlyAlaAlaGlnLys 240
QY 790 CCTGTCACCATTCATGCCACCCAGAGGGGACTTGTGAGCATGGCGCATGATTCTGGA 849
Db 241 ProValThrIleHisAlaThrProGlnGlyThrSerGlnAlaCysArgMetIleLeuGln 260
QY 850 ATCATGCGAAGAGGGGAGATGAGACCAACTAGCCGGAAGAGATTCTGTAAATCTTG 909
Db 261 IleMeGlnLysGlnAlaAspGlnThrLysLeuAlaGlnGlnIleProLeuLysIleLeu 280
QY 910 GCACACATGCGCTTGTGTTGAAGACTGATTGAAAAGAGGACAGAAATTTGAGAAAT 969
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGlnGlyArgAsnLeuLysLysIle 300
QY 970 GAACATGAACAGGACCCAGATACAAATCTCATCTTTCAGAGATTGACATATACAC 1029
Db 301 GlnHisGlnThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleLysTrsn 320
QY 1030 CCGAAGAAACCATCAGCTGTGAAAGGACGACACTTGAAGGCTGTGCGAGCTGAGATAG 1089
Db 321 ProGlnArgThrIleThrValLysGlyThrValGlnAlaCysAlaSerAlaGlnIleGln 340
QY 1090 ATTATGAAGAGCTGCTGAGGCTTGTGAAAATGATATGCTGCGGTGTAACCCACCTCC 1149
Db 341 IleMetLysLysLeuAlaGlnAlaPheGlnAsnAspMetLeuAlaValAsnThrHisSer 360
QY 1150 GGAATCTTCTCAGCTGTACCCCATCACAGTTTGACCCGCTCCGCGATCATCACTCT 1209
Db 361 GlyTyrPheSerSerLeuTyrProHisIleGlnPheGlyProPheProHisHisSer 380
QY 1210 TATCCAGAGAGAGATTGGAATCTTGTATCCCAACCCAGAGCTGTGGGCGCATCATC 1269
Db 381 TyrProGlnGlnGlnIleValAsnLeuPheIleProThrGlnAlaValGlyAlaIleIle 400
QY 1270 GGGAGAGAGGGGGCAGACATCAAAACAGCTGGCGGAGATTGGCGGAGCTCATAGATT 1329
Db 401 GlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIleLysIle 420
QY 1330 GCCCTTCGAGAGGCCAGCGTACGAGAAAGATGGTCATCAACCGGCGCACCGGAA 1389
Db 421 AlaProAlaGlnGlyProAspValSerGlnArgMetValIleIleThrGlyProProGln 440
QY 1390 GCCCAGTTCAAGGCCAGAGGACGAGATCTTTGGGAACTGAAAAGAGAAACTTCTTTAAC 1449
Db 441 AlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGlnGlnAsnPheAsn 460
QY 1450 CCCAAAGAAAGATGAGAGCTGAGAGCGCATATCAGAGTGCCTTTCAGAGCTGAGCGG 1509
Db 461 ProLysGlnGlnValLysLeuGlnAlaHisIleArgValProSerSerThrAlaGlyArg 480
QY 1510 GTGATTGGCAAGAGTGGCAAGACCGTGAAGCAACCTGCAAGAACTTAACAGTGCAGAGTC 1569
Db 481 ValIleGlyLysGlyLysThrValAsnGlnLeuGlnAsnLeuThrSerAlaGlnVal 500
QY 1570 ATGAGCTCTCGTGAACCAAGCGCAGATGAAAATGAGAGAGGTGATCGTCAAGATTATCGGG 1629
Db 501 IleValProArgAspGlnThrProAspGlnAsnGlnGlnValIleValArgIleIleGly 520
QY 1630 CACTTCTTGTAGCCAGACTGACAGCGCCAGAGATCAGGGGAAATTTATACAACAGGTGAAG 1689
Db 521 HisPhePheAlaSerGlnThrAlaGlnAlaGlyLysIleArgGlnIleValGlnGlnValLys 540
QY 1690 CAGCAGACACAAATACCTTCAAGGAGTCCGCTCAACGCGACGAAG 1737
Db 541 GlnGlnGlnGlnLysLysTyrProGlnGlyValAlaSerGlnArgSerLys 556
```

RESULT 3

US-09-764-864-1116

```
; Sequence 1116, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.10
; SEQ ID NO 1116
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (533)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1116

Alignment Scores:
Pred. No.: 1,226-218 Length: 620
Score: 2808.00 Matches: 560
Percent Similarity: 90.78% Conservatve: 1
Best Local Similarity: 90.61% Mismatches: 11
Query Match: 47.71% Indels: 46
DB: Gaps: 2

US-09-270-437d-8 (1-3283) x US-09-764-864-1116 (1-620)
QY 22 GCGCGGCTACCGGGCCGGGGAGCCGGGCTCTCGGGGAGAGACG-----ANG 72
DB 3 ARGAGTGTATACysArgTyrArgSerGlyIleProGlySerThrHisAlaSerGlyMet 22
QY 73 ATGAAGAAGCTTTACATCGGGAACTGAGCGCCGCGTCAACGCGACGACCTCCGGCAG 132
DB 23 MetAsnIysLeuTyrIleGlyAsnLeuSerProAlaIleThrAlaAspLeuArgGln 42
QY 133 CTCTTTGGGAGACAGAAAGCTGCCCTGGCGGACAGGTCCTGCTGAAGTCGGCTACGCC 192
DB 43 LeuPheGlyAspArgIysLeuProLeuAlaGlyGlnValLeuLeuIysSerGlyTyrAla 62
QY 193 TTGCGTACTACCCCGACCAAGCTGGGCTATCCGGCCATCGAGACCCCTCGGGTAAA 252
DB 63 PheValAspTyrProAspGlnAsnTyrPalaIleArgAlaIleGlnThrLeuSerGlyLys 82
QY 253 GTGGAAATTGACGAGAAATCATGAAAGTTGATTACTAGTCTCTAATAAGCTAAGAGC 312
DB 83 ValGlnLeuHisSerGlyLysIleMetGlnValAspTyrSerValSerIysIysLeuArgSer 102
QY 313 AGGAAATTCAGATTGCAAAATCCCTCTCCTACCTGACGAGTGGAGAGTGTGGATGACTT 372
DB 103 ArgLysIleGlnIleArgAsnIleProProHisLeuGlnTyrGlnValLeuAspGlyLeu 122
QY 373 TTGGCTCAATATGGAGACAGTGGAGAAATGTGGAACAATCAACACAGACAGAAACGCC 432
DB 123 LeuAlaGlnTyrGlyThrValGlnAsnValGlnGlnValAsnThrAspThrGlnThrAla 142
QY 433 GTTGTCACGTCACATATGCAACAGAGAGCAAAATATAGCCATGAGAGAACTAAGC 492
DB 143 ValValAsnValIleThrTyrAlaIleThrArgGlnGlnAlaIysIleAlaMetGlnLysLeuSer 162
QY 493 GGGGATCAGTTTGAGAACTACTCTTCAAGATTCTTCAATCCCGATGGAAGAGTAGC 552
DB 163 GlyHisGlnPheGlnAsnTyrSerPheIysIleSerTyrIleProAspGlnGlnValSer 182
QY 553 TCCCTTCGCGCTTCAGCGAGCCCAAGCGGGGAGCACTCTCCCGGAGCAAGGCGCAC 612
DB 183 SerProSerProGlnArgAlaGlnArgIysAspHisSerSerArgGlnGlnIleHis 202
QY 613 GCGCCCTGGGGCACTTCTCAGGCGACAGATGATTCCCGCTGCGGATCCTGCTCCCC 672
```

```
DB 203 AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro 222
QY 673 ACCGAGTTGTTGGTGCATCATCGAAAGAGGGCTTGACCATTAAGACATCACTAAG 732
DB 223 ThrGlnPheValGlyAlaIleIleGlyLysGlnGlyLeuThrIleLysAsnIleThrLys 242
QY 733 CAGACCCAGTCCCGGGTAGATATCCATTAAGAAAGAACTCTGGAGCTCGCAGAGAAAGCCT 792
DB 243 GlnThrGlnSerArgValAspIleHisArgLysGlnAsnSerGlyAlaAlaGlnLysPro 262
QY 793 GTACCATTCATGCGACCCCGAGGGGAGCTCTGAGAGCATGCGCGATGATCTTGAATC 852
DB 263 ValThrIleHisAlaThrProGlnGlyThrSerGlnAlaCysArgMetIleLeuIle 282
QY 853 ATGCAGAAAGAGCAGATAGACCAACTAGCCGGAAGAGATTCTCTGAAAATCTTGCA 912
DB 283 MetGlnLysGlnAlaAspGlnThrLysLeuAlaGlnGlnIleProLeuLysIleLeuAla 302
QY 913 CACAAATGCTTGTTGAAAGACTGATTGGAAAAGAGGACAGAAATTTGAAGAAATTGAA 972
DB 303 HisAsnGlyLeuValGlyArgLeuIleGlyLysGlnGlyArgAsnLeuLysIleGln 322
QY 973 CATGAAGACAGGACCAAGATTAACATCTCATCTTGCAGATTGAGCATATACACCG 1032
DB 323 HisGlnThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 342
QY 1033 GAAAGAACCATCATCTGTGAAGGGCAGCTTGAAGGCTTGCCAGTGTGATGATGAGATT 1092
DB 343 GlnArgThrIleThrValLysGlyThrValGlnAlaCysAlaSerAlaGlnIleGlnIle 362
QY 1093 ATGAAGAAAGCTGCGTGGAGGCTTTGAAAATGATATAGCTGCGTGAAC----- 1140
DB 363 MetLysLysLeuArgGlnAlaPheGlnAsnAspMetLeuAlaValAsnGlnAlaAsn 382
QY 1140 ----- 1140
DB 383 LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 402
QY 1141 -----ACC 1143
DB 403 SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThrThr 422
QY 1144 CACTCCGAAATCTTCTCCAGCTGTACCCCATCACAGATTGGCGCGTCCCGCATCAT 1203
DB 423 HisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHis 442
QY 1204 CACTCTTATCCAGAGAGAGGATTTGAAATCTCTGATCCCAACCGAGCTGGGGGCC 1263
DB 443 HisSerTyrProGlnGlnGlnIleValAsnLeuPheIleProThrGlnAlaValGlyAla 462
QY 1264 ATCATCGGAGAAAGAGGGGACACATCAACACAGCTGGCGAGATTCCCGGAGCTCTATC 1323
DB 463 IleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle 482
QY 1324 AAGATTGCCCTTCGGAAGGCCCGACAGCTGAGGAAAGATGTGCATCATCACCGGSCCA 1383
DB 483 LysIleAlaProAlaGlnGlyProAspValSerGlnArgMetValIleIleThrGlyPro 502
QY 1384 CCGGAGGCCAGTTCAAGGCCCGAGGAGCGATCTTTGGGAACTGAAAGAGAAACTTC 1443
DB 503 ProGlnAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGlnGlnAsnPhe 522
QY 1444 TTTAACCACCAAGAAAGAGTGAAGCTGGAAGCGCATTCAGAGTCCCTCTTCCACAGCT 1503
DB 523 PheAsnProLysGlnGlnValLysLeuGln**HisIleArgValProSerSerThrAla 542
QY 1504 GCGCGGCTGATTGCAAGAGTGGCAAGCTGCAAGCGTGAACGATCGCAAGCTTACAGTGA 1563
DB 543 GlyArgValIleGlyLysGlyLysGlyLysThrValAsnGlnLeuGlnAsnThrSerAla 562
QY 1564 GAAATCATCGCTCGTGCACCAAGCGCAGATGAAATGAGAAATGATGTCGTCGAAATT 1623
DB 563 GlnValIleValProArgAspGlnThrProAspGlnAsnGlnGlnValIleValArgIle 582
```


QY 1504 GCGCGGCTGATGTGCAAGGTGGCAAGACCGTGAACGAACTTACCACTTACCAAGTCA 1563
Db 521 G1YArgVal1IleGly1ysGly1ysThrValAsnGluLeuGlnAsnLeuThrSerAla 540
QY 1564 GAAGTCATCCGCTCGTGCACCAAGCCAGATGAAATGAGAAATGATCGTCAAAAT 1623
Db 541 G1Val1IleVal1Pro1Arg1Asp1Gln1Thr1Pro1Asp1Glu1Asn1Glu1Val1IleVal1Arg1Ile 560
QY 1624 ATCCGGGACCTCTTGTGTAGCCAGACTGCGACAGCGCAAGATCGAGAAATGTTCACACAG 1683
Db 561 I1leG1His1Phe1Phe1Ala1Ser1Gln1Thr1Ala1Gln1Arg1Lys1Ile1Arg1Glu1IleVal1Gln1Gln 580
QY 1684 GTGAAGCAGCGACGACAGCAAAA 1704
Db 581 Val1ysGln1Gln1Gln1Gln1Lys 587

RESULT 5
US-10-262-445-40
Sequence 40, Application US/10262445
Publication No. US20040014058A1
GENERAL INFORMATION:
APPLICANT: Alsobrook II, John
APPLICANT: Burgess, Catherine
APPLICANT: Catereron, Elina
APPLICANT: Chant, John
APPLICANT: Chaudhuri, Amitabha
APPLICANT: Edinger, Shlomit
APPLICANT: Gerlach, Valerie
APPLICANT: Giot, Loic
APPLICANT: Gorman, Linda
APPLICANT: Guo, Xiaojia
APPLICANT: Kekuda, Ramesh
APPLICANT: Mezes, Peter
APPLICANT: Millet, Isabelle
APPLICANT: Ooi, Chean Eng
APPLICANT: Paturajan, Meera
APPLICANT: Rieger, Daniel
APPLICANT: Spytek, Kimberly
APPLICANT: Taupier Jr., Raymond J.
APPLICANT: Zerhusen, Bryan
APPLICANT: Zhong, Haihong
APPLICANT: Zhong, Mei
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
FILE REFERENCE: 21402-462D
CURRENT APPLICATION NUMBER: US/10/262,445
CURRENT FILING DATE: 2002-10-01
PRIOR APPLICATION NUMBER: 60/327,454
PRIOR FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 60/327,917
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/328,029
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/328,056
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/328,849
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/329,414
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 60/330,142
PRIOR FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: 60/341,058
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 60/343,629
PRIOR FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 60/349,575
PRIOR FILING DATE: 2001-10-29
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 133
SOFTWARE: Curseseq version 0.1
SEQ ID NO: 40
LENGTH: 555
TYPE: PRT

ORGANISM: Homo sapiens
US-10-262-445-40
Alignment Scores:
Pred. No.: 1,61e-200 Length: 555
Score: 2584.50 Matches: 507
Percent Similarity: 94.60% Conservative: 19
Best Local Similarity: 91.19% Mismatches: 29
Query Match: 43.91% Indels: 1
Gaps: 1
US-09-270-437d-8 (1-3283) x US-10-262-445-40 (1-555)
QY 70 ATGATGAACAAGCTTTATCATGCGGAACTGAGCCCGCCGCTCAGCCGAGACCTCGG 129
Db 1 MetMetAsnLys1LeuPhe1IleG1YAsn1LeuSerPro1Ala1Val1Ala1Glu1Spu1Arg 20
QY 130 CAGCTCTTTGGGACAGGAAAGTGGCCCTGGCGGAGACAGTCTGCTGAAGTCCGCTAC 189
Db 21 Gln1euPhe1Gly1Asp1Arg1Lys1Leu1Pro1Leu1Ala1Gly1Gln1Val1Leu1Lys1Ser1Arg1Tyr 40
QY 190 GCCTTCGTGACTACCCCGACAGAACTGGGCGCATCGCGCCATGAGACCTCTCGGCT 249
Db 41 Ala1Phe1Val1Asp1Tyr1Pro1Asp1Gln1Ser1Tyr1Ala1Ile1Arg1Thr1Ile1Glu1Thr1Leu1Ser1Gly 60
QY 250 AAAGTGAATTGCGATGGGAAATATCATGGAAGTGAATTAATCTGAGTCTGCTAAAGGTAAG 309
Db 61 Gln1Val1Glu1Leu1His1Gly1Lys1Ile1Met1Glu1Val1Asp1Tyr1Ser1Val1Ser1Lys1Leu1Arg 80
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCTCACTGCGACAGTGGAGGAGTGTGATGGA 369
Db 81 Ser1Arg1Asn1Ile1Pro1Ile1Arg1Asn1Ile1Pro1Ile1Ser1Gln1Ile1Thr1Glu1Val1Leu1Asp1Gly 100
QY 370 CTTTGGCTCATATATGAGACAGTGGAAATGTGAAACAGTGAACACACACACACAGAAAC 429
Db 101 Leu1eu1Ala1Gln1Tyr1Gly1Thr1Val1Glu1Asn1Val1Glu1Gln1Val1Asn1Thr1Arg1Thr1Glu1Thr 120
QY 430 GCCGTTGTACAGTCATATATGCAACAGAAAGAAAGAAAGAAATAGCCATGGAGAGCTA 489
Db 121 Ala1Val1Asn1Val1Thr1Tyr1Ala1Thr1Lys1Glu1Gln1Val1Lys1Ile1Ala1Met1Lys1Lys1Leu 140
QY 490 AGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTCTTACATCCCGATGGAAGAGGTG 549
Db 141 Ser1Gly1His1Gln1Phe1Glu1Asn1His1Tyr1Phe1Lys1Ile1Ser1Tyr1Ile1Pro1Asp1Arg1Val1 160
QY 550 AGTCCCTCTTGGCCCTCAGCGAGCCAGACGCTGGGACCACTTTCCCGGAGGACAGGC 609
Db 161 Ser1Cys1Pro1Ser1Pro1Gln1Arg1Ala1Gln1Arg1Lys1Asp1His1Ser1Ser1Tyr1Glu1Gln1Gly 180
QY 610 CACGCCCCCTGGGGGCACTTCTCAGGCCAGACAGATTTCCCGCTGCGGATCTGTGC 669
Db 181 Gln1Ala1Pro1Gly1Lys1Ser1Ser1Gln1Ala1Arg1Gln1Ile1Asp1Phe1Pro1Leu1Ala1Glu1Phe 200
QY 670 CCCAGCCAGTTTGTTGGTGCATCATCGGAAAGAGGCTTGAACATTAAGAACATCACT 729
Db 201 Pro1Thr1Gln1Phe1Val1Gly1Ala1Ile1Ile1Gly1Lys1Glu1Gly1Leu1Thr1Ile1Lys1Asn1Ile1Thr 220
QY 730 AAGCAGACCCAGTCCCGGATGATATTCATTAAGAAAGAAAGAACTTGGAGCTGCGAGAGAAG 789
Db 221 Lys1Gln1Ser1Arg1Ser1Arg1Val1Asp1Ile1Tyr1Arg1Gln1Gln1Asn1Ser1Arg1Ala1Ile1Glu1Lys 240
QY 790 CTTGTACACATCCATGCCACCCCGAGAGGAGCTTCTGAAGCATCCGCGATCTTGA 849
Db 241 Pro1Val1Thr1Met1His1Ala1Thr1Pro1Glu1Gly1Thr1Ser1Glu1Ala1Cys1Arg1Met1Ile1Leu1Glu 260
QY 850 ATCATGCAAGAAAGGAGGAGATGATGACCAACTAGCCGAGAGATCTCTGAAATCTTG 909
Db 261 Ile1Met1Gln1Lys1Glu1Ala1Asp1Glu1Ala1Lys1Leu1Ala1Glu1Glu1Ile1Pro1Leu1Lys1Ile1Leu 280
QY 910 GCACACATGCGCTTGTGTGAAGACTGATTGAAAGAGGACGACAAATTTGAAGAAATT 969
Db 281 Ala1His1Asn1Gly1Leu1Val1Gly1Arg1Leu1Ile1Gly1Lys1Glu1Gly1Arg1Asn1Leu1Lys1Lys1Asn 300

```

QY      970 GAACATGAAACAGGGGACCAAGATACATCTCATCTTTCAGAGATTGAGCATATACAAAC 1029
Db      301 GlnHsGslThrGlyThrIysIleThrIleSerSerSerGlnAspSerIleIyrAsn 320
QY      1030 CCGGAAAGAACCATCATCTCTGAAAGGGACACAGTTGAGCCTGTGCCAGTCTGAGATAGAG 1089
Db      321 ProGlnArgThrIleThrValIysGlyThrValGluValCysAlaSerIleGluIleGln 340
QY      1090 ATTATGAAGAGCTGGCTGAGGCTTGTGAAATGATATGCTGGCTTTAAACCCACCTCC 1149
Db      341 IleMetIlybLysLeuHrGlnIleValAsnLeuPheIleProThrGlnIlyValGlyAlaIleIle 360
QY      1150 GGAATCTTCTCCAGCCTGTACCCCATCACAGTTTGAGCCCTGCCGATCATCATCTCT 1209
Db      361 GlyTyrPheSerSerIleuIyrProHlsArgGlnPheGlyProPheProHlsHsSer 380
QY      1210 TATCCAGAGCAGGAATGATGTGATCTTCTTATCCCAACCCAGGCTGTGGGCCCATCATC 1269
Db      381 TyrProGlnGlnGlnIleValAsnLeuPheIleProThrGlnIlyValGlyAlaIleIle 400
QY      1270 GGGAGAGAGGGGGCACATCATCAACAGCTGGCGGATTCGCCGGAGCTTATCAAGTT 1329
Db      401 GlyIysIysGlyAlaHlsIleIysGlnLeuAlaArgPheValGlyAlaSerIleIysIle 420
QY      1330 GCCCTGCGGAAAGGCCACGCTACGCGAAAGATGATGATCATCATCACCGGGCCACCGGAA 1389
Db      421 AlaProAlaArgSerPro--LeuHrGlnArgIysValIleIleThrTyrProProGln 439
QY      1390 GCCCATGTCAGAGCCCAAGGACGATCTTTGGGAAACTGAAAGAGAAACTTCTTTAAC 1449
Db      440 SerGlnPheIysAlaGlnIlyArgIlePheGlyIysLeuIysGlnIlyAsnPheHsAsn 459
QY      1450 CCCAAGAGAGAGTGAAGCTGGAAGCGCATATCCAGAGGCGCTTCCACAGCTGGCCGG 1509
Db      460 ProIysGlnAspValIysLeuGlnIthrHlsIleArgValProSerSerThrAlaGlyArg 479
QY      1510 GTGATTGGCAAGGTGGGACAGCCGTGAAACGACTGCGAAGCTTAACAGTGCAGAGATC 1569
Db      480 ValIleGlyIysGlyIysThrValAsnGlnLeuGlnAsnLeuIleSerAlaGluVal 499
QY      1570 ATCGTGCCTGCTGACCAAAAGCCGATGATGAATGAGAGATGATGTCAGATTATCGGG 1629
Db      500 IleValProArgAspGlnIthrProAspGlnAsnGlnGluMetIleValArgIleIleGly 519
QY      1630 CACTCTTGTGCTACCGAGACTGCACAGGGGCAAGATCGGGAATTTGTAACACAGTGAAG 1689
Db      520 HisPhePheAlaSerGlnIthrAlaGlnArgIysIleArgGluIleValGlnGlnValIys 539
QY      1690 CAGCAGAGAGCAAAATACCTTCAGGAGTGCCTCAGAGCGCAGCAAG 1737
Db      540 GlnGlnGlnGlnIlyIysTyrProGlnGlnIlyValAlaSerGlnArgSerIys 555

RESULT 6
US-10-313-986-500
; Sequence 500, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: McNabb, Andrea M.
; APPLICANT: Foy, Teresa M.
; APPLICANT: McManabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 577
; TYPE: PR
; ORGANISM: Homo sapiens

```

```

US-10-313-986-500
Alignment Scores:
Pred. No.: 3,34e-148 Length: 577
Score: 1940.00 Matches: 383
Percent Similarity: 79.10% Conservative: 75
Best Local Similarity: 66.15% Mismatches: 95
Query Match: 32.96% Indels: 26
DB: 15 Gaps: 5

US-09-270-437d-8 (1-3283) x US-10-313-986-500 (1-577)
QY      73 ATGAACAAGCTTTAATCGGAGAACTGAGCCCGCGTCAACCGGACGACCTCCGCGAG 132
Db      1 MetAsnIlybLeuIyrIleGlyAsnLeuAsnGlnSerValIthrProAlaAspGlnIlyAs 20
QY      133 CTCTTGGGGAACAGGAACTGCCCCCTGGCCGGACAGCTCTCTGTAAGTCCGGCTTAACGCC 192
Db      21 ValPheAlaGlnHlsIlybLeuIysSerIysGlnPheLeuValIysSerIlyIyrAla 40
QY      193 TTGCTGACTACCCCGACAGACTGGGCAATCCGCGCATGAGACCTCTCGGGTAA 252
Db      41 PheValAspCysProAspGlnHlsIthrAlaMetIysAlaIleGlnIthrPheSerGlyIys 60
QY      253 GTGCAATTGCAATGGGAAATCATGGAAGTTATTACTCAGTCTTAAAGAGCTAAGAGC 312
Db      61 ValGlnLeuGlnIlyIysArgLeuGlnIleGlnHlsSerValProIlybGlnArgSer 80
QY      313 AGGAAATTCAGATTTCGAAACATCCCTCTCAGCTGAGGTGGAGGTGTGATGACTT 372
Db      81 ArgIlybIleGlnIleArgAsnIleProProGlnLeuArgIyrGlnValLeuAspSerLeu 100
QY      373 TTGCTCATATGAGGACAGTGGAGATGTGGAACAAGTCAACACAGACAGCAAGACCGGC 432
Db      101 LeuAlaGlnIyrGlyIthrValGlnAsnCysGlnGlnValAsnIthrGlnSerGlnIthrAla 120
QY      433 GTTGTCAAGCTCATATGTCAACAAGAGAGACCAAAAATAGCCATGAGAGAGCTAAGC 492
Db      121 ValValAsnValThrIyrSerAsnArgIleGlnIthrArgIleAlaIleMetIlybLeuAsn 140
QY      493 GGGCATGATTGGAAGTCTCTCTTCAAGATTCCTACATCCCGGATGAGAGAGGTGAGC 552
Db      141 GlnIysGlnLeuGlnAsnHlsAlaLeuIysValSerIyrIleProAspGlnIleAla 160
QY      553 TCCCTTCGCCCCCTCCAGGACGACGCTGGGACCACTCTTCCCGGAGCAAGGCCAC 612
Db      161 -----GlnGlySerProGlnAsnGlnArgArgIlyIysIleGlySerArgIlyGlnProArg 178
QY      613 -----GCCCTGGGGGCACTTCTCAGGCCACAGATGATTGTTCCCGTG 657
Db      179 GlnGlySerProValAlaIleAlaIlyAlaProAlaIlybGlnGlnIlyAspIleProLeu 198
QY      658 CGGATCGTGTCCCGACCCGATTTGTTGTGCCATCATCGGAAAGAGGGCTTGACATA 717
Db      199 ArgLeuLeuValProIthrGlnIyrValGlyAlaIleIleGlyIysGlnIlyAlaIthrIle 218
QY      718 AAGAACATCACTAAGAGACCCAGTCCCGGTGATTCATTCATGAAAGAACTCTGGA 777
Db      219 ArgAsnIleThrIysGlnIthrGlnSerIysIleAspValHlsArgIysGlnAsnAlaGly 238
QY      778 GCTGCAGAGAGCGCTGACCATCATGCGACCCCAAGGGGCACTTCTGAAGATGCGGC 837
Db      239 AlaAlaGlnIlybAlaIleSerValHlsSerThrProGlnIlyCysSerAlaCysIys 258
QY      838 ATGATCTTGAATTCATGCAAGAGGAGGAGATGAGACCAACTAGGCCGAAGATTCCT 897
Db      259 MetIleLeuGlnIleIleHlsIlybGlnIlyAspIlybIlybIlybIlybIlybIlybIlyb 278
QY      898 CTGAAATCTTGGCACACATGCTGTGGTGAAGCTGATTGGAAGAGAGGAGCAAT 957
Db      279 LeuIysIleLeuAlaHlsAsnAsnPheValGlyArgLeuIleIlyIysGlnIlyArgAsn 298
QY      958 TTGAAGAAATTCAGATGAAACAGGGACCAAGATTAACAATCTCATCTTTCAGAGATTGG 1017

```

[illegible]

Alignment Scores:	1,03e-147	Length:	577
Pred. No.:	1934.00	Matches:	381
Score:	79.10%	Conservative:	77
Percent Similarity:	65.80%	Mismatches:	95
Best Local Similarity:	32.86%	Indels:	26
Query Match:		Gaps:	5
DB:	9		
US-09-270-437D-8 (1-3283) x US-09-873-6337-2 (1-577)			
QY	73 ATGAACAAGCTTTACATCGGGAACTCGAGCCCGCGTACCGCCAGACCGACTCCGGCAG	132	
Db	1 MetasnysleuYrllleGlyasneLueasGluSeValThrProAlaspneuGluys 20		
QY	133 CTCCTTGGGGACAGGAAGCTGCCCCCTGGGGGACAGCTTCCTGATCGACGCTACCGC	192	
Db	21 ValPhealagluHisIuylsIleSeTyYrseGdlGlnPheLueValIysSeGdlYrYrAla 40		
QY	193 TTGCTGACATACCCCGACCGAACTGGGCCATCCGGCCATGAGAACCTCTCGGGTAA	252	
Db	41 PheValaspCySProAspGluHisIeTrpAlaMetYrYrAlaIleGluThrPheSeGdlYys 60		
QY	253 GTGAAATTGCATGGGAAATCATGGAAGTTACTCATGCTCTCTAAAGCTAAAGGAC	312	
Db	61 ValGluLeuGlnGlyIysArgLeuGluMetGluHisSeValProIlysluSglnaGsser 80		
QY	313 AGGAAAATTGAGTTTCGAAACATCCCTCCACCTCAGTGGAGGAGGTGGATGACTT	372	
Db	81 ArgYslIleGlnIleHagSnIleProProGlnLeuArgTrpGluValLeuaspSeIleu 100		
QY	373 TTGGCTCAATATGGGACAGTGGAGATGTGGAAACAGTCAACACACACACAGAACCGGC	432	
Db	101 LeuAlaGlnYrYrGlyThrValGluAsnCyGluGlnIleValAsnHrgIuSeGdluThrAla 120		
QY	433 GTTGTCAACGTCAATATGCAACAAGGAGGAAGCAAAATAGCCATGGAGAACTAAGC	492	
Db	121 ValValaenValThrYrSerAsnHrgIuGlnHrArgGlnAlaIleMetYsluEasH 140		
QY	493 GGGCATCGTTTGAAGAACTACTCCCTCAAGATTTCTTACATCCCGAGTAAGAGGTGAC	552	
Db	141 GlyHsGlnLeuGluAsnHsAlaLeuHysValSeTyYrIleProAspGluGlnIleThr 160		
QY	553 TCCCTTGGCCCCCTCAGCGAGCCCGAGCTGGGGACACTCTTCCCGGAGCAAGGCCAC	612	
Db	161 -----GlnGlyProGluAsnGlyArgHrgGlyYrGlyPheGlySeIArgYslGlnProArg 178		
QY	613 -----GCCCTGGGGGGCATCTTCAGCGACGACAGATTGATTTCGCCGCTG	657	
Db	179 GlnGlySerProValAlaAlaGlyAlaProAlaIuSglnGlnProValaapIleProLeu 198		
QY	658 CGGATCGTGGTCCCGACCCAGTTGTTGTGTCATCATCGGAAAGAGAGGCTTGACATA	717	
Db	199 ArgLeuLeuValProThrdlnYrValGlyAlaIleIleGlyIuSglnGlyAlaThrIle 218		
QY	718 AAGAACATCACTAAGCAACCCAGTCCGGGTAGATTCCTATAGAAAAGAACTTCGGA	777	
Db	219 ArgAsnIleThrIuSglnHrHrgInSeTyIleAspValHsArgIuSgluAsnAlaGly 238		
QY	778 GCTGCAGAGAAGCCGTGCATCATCGATCGACCCGACAGGGGACCTTCTGAAGCATGCGC	837	
Db	239 AlaAlaGluHysAlaIleSeValHsSerThrProGluGlyCySseSeSerAlaCyLys 258		
QY	838 ATGATCTTGAATCATGACGAAGAAGGACAGATGAGACCAACTAGCCGGAAGATTCCT	897	
Db	259 MetIleIeuGlnIleHethrIuSyluGlnAlaIuYsAspThrIuYsTrnAlaAspGluValPro 278		
QY	898 CTGAAATCTTGGCACACAATGGCTTGTTGGAAAGACATGATTGGAAAAGAGGACGAAT	957	
Db	279 LeuYrIleLeuAlaHisAsnHsPheValGlyArgLeuIleGlyIuSglnGlyYrArgSn 298		
QY	958 TTGAAGAAATTTGAACATGAACAGGAGCAAGATTAACATCTCATCTTTCAGGATTGG	1017	


```

QY 961 AAGAAATGTAACATGAAACGGGACCAAGATTAATCATCTTTGGAGATTGAGC 1020
Db 300 LysylsilegyluglnaspThrAspThrileThrileSerProleghleuThr 319
QY 1021 ATATACAAACCGGAAAGAACCATCTGTAAGGGGACAGATTGAGCCTGTGCCAGTCT 1080
Db 320 LeuylAsnProgluArgThrileThrValylsGlyAsnValgluThrCysAlaIysAla 339
QY 1081 GAGATGAGATTATGAAAGAGCTGCGTAGGCGCTTTGAAATGATATGCTGCTTAA 1140
Db 340 GluGluGluIleMetLysylleArgIleuSerYrGluAsnAspIleAlaSerMetAsn 359
QY 1141 ACCCACTCC-----GATACCTC----- 1158
Db 360 LeuGlnAlaHslleuIleProgluLeuAsnAlaAlaLeuGluLeuPheProthr 379
QY 1159 -----TCCAGCTGTACCCCATCAACAGTTTGCCCTTCCCGCAT 1200
Db 380 SerGlyMetProProProThrSerGlyProProSerAlaMetThrProProIyProGln 399
QY 1201 CATCACTCTTATCCAGACGAGAGATTGAAATCTTTCATCCCAACCGCTGTGGGC 1260
Db 400 PhegluGln---SerGluThrGluThrValHslleuPheIleProAlaLeuSerValGly 418
QY 1261 GCCATCATGCGGAAGAGGGGCAACATCAACAGCTGCGGAGATTGCGCGAGCCTCT 1320
Db 419 AlaIleIleGlyLysglInhslsileLysGlnLeuSerArgPheAlaGlyAlaSer 438
QY 1321 ATCAAGATTGCCCGCGGAGAGCGCCAGACGTCAGCGAAAGATGCTCATCATCCCGG 1380
Db 439 IleLysIleAlaProAlaGluAlaProAlaAlaLysValArgMetValIleIleThrGly 458
QY 1381 CCACCGGAACCCCGATTCAAGGCCCAAGGACGATCTTTGGGAACTGAAAGGAAAC 1440
Db 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrglyLysIleLysGluGluAsn 478
QY 1441 TTTCTTAACCCCAAGAAAGAGTGAAGCTGGAAGCCATATCAGAGTGCCTTTCCACA 1500
Db 479 PheValSerProLysGluGluValLysLeuGluAlaHslleArgValProSerPheAla 498
QY 1501 GTCGCGCGGATGATGCAAGAGTGCGCAAGACCGTGAACGAACTTGCAAACTTAACCA 1560
Db 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuSer 518
QY 1561 GCAGAAATCATGCTGCTGCTGTAACCAACCGCCAGATGAATAAGAAAGATGCTGCA 1620
Db 519 AlaGluValAlaValProArgAspGlnThrProAspIleAsnAspGlnValAlaLys 538
QY 1621 ATATCGGGGACCTTTGTGTAAGCACTGCAAGCGGCAAGATCAGGAAATTTTACA 1680
Db 539 IleThrIleHisPheIlyAlaCysGlnValAlaGlnIleArgLysIleGlnGluIleLeuThr 558
QY 1681 CAGGTGAAGAG---CAGGAGCAGAAATACCTTCAG---GAGTGCCTTCACAGCGCAGC 1734
Db 559 GlnValLysGlnHslsGlnGlnGlnIleAlaLeuGlnSerGlyProProGlnSerArgArg 578
QY 1735 AAGTGAAGCTCCCAAGGACACCGCAAAACAC 1767
Db 579 LysHslsHslsHslsHslsHslsHsls 589

```

```

; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C14
; CURRENT FILING DATE: US/09/735.705
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-705-348

Alignment Scores:
Pred. No.: 6,32e-143 Length: 579
Score: 1875.00 Matches: 369
Percent Similarity: 78.70% Conservative: 78
Best Local Similarity: 64.96% Mismatches: 95
Query Match: 31.86% Indels: 26
Gaps: 6
DB: 9

US-09-270-437d-8 (1-3283) x US-09-735-705-348 (1-579)
QY 73 ATGAAACAAGCTTTATACATGCGGAACCTGAGCCCGCGCTGACCGCGACGACTCCGGCAG 132
Db 1 MetAsnLysLeuIlyIleGlyAsnLeuSerGluAsnAlaAlaProSerAspIleuSer 20
QY 133 CTCTTTGGGAGAGGAAGCTGCGCTGCGGAGACAGTCTGCTGAAAGTCGGCTACGCGC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyra 40
QY 193 TTCGTGACTACCCGACACAGAACTGGGGCCATCCGGCCATGAGACGCTTGGGTAA 252
Db 41 PheValAspCysProAspIleuSerIleProAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAATTCATGCGGAATTCATGAAAGTGTATCTCACTCACTCTTAAAGCTTAAGAGC 312
Db 61 IleGluLeuHslsGlyLysProIleGluValGlnHslsSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCATGATTCGAAACATCTCTCTACCTGCACTGCGGAGGTGTGATGACT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHslsLeuGlnIleProAlaLeuAspSerLeu 100
QY 373 TTGGCTCATATGCGGACAGCGGAAATGTGAAACAAGTCAACACAGACAGAAACCGCC 432
Db 101 LeuValGlnIlyGlyAlaValGluSerCysGlnGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGCAACGTCATATGCAACAGAAAGAAAGCAAAATAGCCATGAGAACTTAAGC 492
Db 121 ValValAsnValThrIlySerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCAATCATTGAAGACTACTCTTCAAGATTTCCACATCCCGGATGAAGAGTAGC 552
Db 141 GlyPheGlnLeuGlnAsnPheThrLeuLysValAlaIlyIleProAspGlnThrAlaAla 160
QY 553 TCCCTTCGCGCCCTTCAG-----CGAGCCAGCGT-----GGGACCACTTTCGCG 600
Db 161 GlnGlnAsnProLeuGlnGlnProAlaGlyLysArgLysGlnGlnGlnGlnGlnGlnGln 180
QY 601 GAGCAAGGACCGCGCTGCGGACATCTTCAGGCGACAGACAGATGATTTCCGCTGCGG 660
Db 181 ArgGlnIly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGCTCCCAACCGCTTGTGTGTCATCATCGGAAGAGGCTTGAACCTAAG 720
Db 200 LeuLeuValProThrGlnPheValAlaIleIleGlyLysGlnGlyAlaThrIleArg 219
QY 721 AACATCACTAAGACAGACCCCACTCCCGGCTAGATTCATAGAAAAGAACTTGGAGCT 780

```

Db 220 AsnIlethrIysgInthrInserIlysIleAspValHisArgIysgInAsnIleAlaGlyAla 239
QY 781 GCAAGAAAGCTGTACCATTCATGCCACCCAGAGGGAGCTTGTGAAGATGCCGATG 840
Db 240 AlaGlnIysSerIleThrIleLeuSerThrProGlnIlyThrSerIleAlaIleCysIysSer 259
QY 841 ATTCTGAATCTATGCAAGAAAGGAGGATGAGAACCAATACCCAGAAAGATTCCTCTG 900
Db 260 IleLeuGlnIleMetHisIleGlyIleAlaGlnIleIleIysPheThrGlnIleProLeu 279
QY 901 AAAATCTGSCACAAATGCTGTGAAGACTGATGAAAGAAAGAGGACAGAAATTTG 960
Db 280 LysIleLeuAlaHisAsnAsnIlePheValGlyArgLeuIleGlyIleGlnIleValGlnIle 299
QY 961 AAGAAATTTGACATGAAACAGAGGACCAAGATTAACAATCTCATCTTTGACGATTTGAGC 1020
Db 300 LysIleIleGlnIleGlnIleAspThrIleAspThrIleIleThrIleSerProLeuGlnIleThr 319
QY 1021 ATTATCAACCCGGAAGAAACCATCATCTGTGAAGGACAGTGAAGGCTGTGCCAGTGTCT 1080
Db 320 LeuIleIleAsnProGlnIleIleThrIleThrValIleIleGlyIleAsnValGlnIleCysAlaIle 339
QY 1081 GAGATTAAGATTTATGAAGAGCTGCGTGAAGGCTTTGAAATGATATGCTGCTGTTAAC 1140
Db 340 GlnGlnIleIleMetIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 359
QY 1141 ACCCACTTC-----GGATTCCTC----- 1158
Db 360 LeuGlnIleAlaHisIleIleIleProGlyLeuAsnIleAsnAlaIleGlyIlePheProIleThr 379
QY 1159 -----TCCAGCTGTACCCCATCAACCATGTTGGCCGCTTCCGCGAT 1200
Db 380 SerGlyMetProIleProIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 399
QY 1201 CATCATCTTATTCAGAGGAGAGATTTGATGATCTCTTCAATCCCAACCCAGGCTGTGGC 1260
Db 400 PheGlnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 418
QY 1261 GCCATCATCGGAGAGAGAGGAGGAGCACAATCAACCAAGCTGCGAGATTCGCGGAGCTCT 1320
Db 419 AlaIleIleIleGlyIleGlnIleGlnIleIleIleIleIleIleIleIleIleIleIleIleIle 438
QY 1321 ATCAAGATTTGCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db 439 IleIleIleIleAlaProIleAlaIleIleIleIleIleIleIleIleIleIleIleIleIleIle 458
QY 1381 CCACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 459 ProIleGlnIleAlaGlnIlePheIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 478
QY 1441 TTCTTTAACCCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Db 479 PheIle 498
QY 1501 GCTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Db 499 AlaGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 518
QY 1561 GCAGAACTCATCGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
Db 519 AlaIle 538
QY 1621 ATTATCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
Db 539 Ile 558
QY 1681 CAGGTGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1704
Db 559 GlnValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 566
US-09-850-716A-348

Sequence 348, Application US/09850716A
Patent No. US20020115139A1
GENERAL INFORMATION:
APPLICANT: Kalos, Michael D.
APPLICANT: McNeill, Patricia D.
APPLICANT: Retter, Marc W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C15
CURRENT APPLICATION NUMBER: US/09/850, 716A
NUMBER OF SEQ ID NOS: 440
CURRENT FILING DATE: 2001-05-07
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 348
LENGTH: 579
TYPE: PRT
ORGANISM: Homo sapiens
US-09-850-716A-348

Alignment Scores:

Pred. No.:	6,32e-143	Length:	579
Score:	1875.00	Matches:	369
Percent Similarity:	78.70%	Conservative:	78
Best Local Similarity:	64.96%	Mismatches:	95
Query Match:	31.86%	Indels:	26
DB:	9	Gaps:	6

US-09-270-437D-8 (1-3283) x US-09-850-716A-348 (1-579)

QY 73 ATGACAAAGCTTATCATTCGGAAGCTGAGCCCGCGCTCACCGCGAGCATTCGCGAGC 132
Db 1 MetAsnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 132
QY 133 CTCTTTGGGAGCAGAAAGCTGCGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 192
Db 21 IlePheIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 40
QY 193 TTGCTGACCTACCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 252
Db 41 PheValAspIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 60
QY 253 GTGGAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 312
Db 61 IleGlnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 80
QY 313 AGAAATTTGAGTTCGAAATCATCT 372
Db 81 ArgIle 100
QY 373 TTGCTGATATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 432
Db 101 LeuValGlnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 120
QY 433 GTTGTCAAGCTCATATTCGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 492
Db 121 ValValAsnIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 140
QY 493 GGGCATGATTTGAGAACTACTCTCTCAAGATTTCTTACATCCCGGATGAAGAGGAGGAG 552
Db 141 GlyPheIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 160
QY 553 TCCCTTCGCGCCCTCG-----CGAGCCAGGCT-----GGGAGCACTCTTCCCGG 600
Db 161 GlnIle 180
QY 601 GAGCAAGCAGCAGCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 181 ArgIle 199
QY 661 ATCTGTGCTCCCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720
Db 200 LeuIle 219

```

QY 721 AACATCACTAAGCAGACCCAGTCCCGGATATATCCATAGAAAAGAACTCTGGAGT 780
Db 220 AsnIleThrIysGlnThrGlnSerIleAspValHisArgIleGlnHisAlaIleVal 239
QY 761 GCGAGGAAGCTGTACCATCCATCCACCCAGAGGGAGCTTGTGAAGATGCGCATG 840
Db 240 AlaGlnIysSerIleThrIleLeuSerThrProGlnGlyThrSerAlaIleCysIysSer 259
QY 841 ATTCTTGAATCATGACAGAAAGAGAGTGAACCACTGACCGAAGAGTTCCTCTG 900
Db 260 IleLeuGlnIleMetHisLysGlnValIleGlnMetIleLysPheThrGlnGlnIleProLeu 279
QY 901 AAAATCTTGGCAACACATGGCTTGGTGGAGACTGATTGGAAAAAGAGCAGAAATTG 960
Db 280 LysIleLeuAlaHisAsnAsnMetValGlyArgLeuIleGlyLysGlnIleYargSerLeu 299
QY 961 AAGAAATTTGAACATGAAACAGGAGCCAAAGATTAACATCTCATCTTTGAGAGATTGAGC 1020
Db 300 LysLysIleGlnGlnAspThrAspThrLysIleThrIleSerProLeuGlnGlnIleThr 319
QY 1021 ATATACAAACCCGAAAGAACCATCACTGTGAAAGGGCAGACATTGAGCGCTGTGCT 1080
Db 320 LeuTyrsAsnProGlnIleThrIleThrValLysGlnValGlnThrCysAlaLysAla 339
QY 1081 GAGATGAGATTATGAAAGAGCTGCGTGAAGGCTTTGAAATGATATGCTGCTGTATAC 1140
Db 340 GlnGlnGlnIleMetLysIleArgIleGlnSerTyrglnAsnAspIleAlaSerMetAsn 359
QY 1141 ACCCACTCC-----GATCTTC----- 1158
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuMetAlaLeuGlyLeuPheProThr 379
QY 1159 -----TCCAGCTGTACCCCGCATCACAGTTTGGCCGTCGCCGAT 1200
Db 380 SerGlyMetProProProThrSerIleProProSerAlaMetThrProProIleProGln 399
QY 1201 CATCACTCTTATCCAGACGAGAGATTGTGAATCTTTCATCCCAACCCAGCTGTGGGC 1260
Db 400 PheGlnGln-----SerGlnThrGlnThrValHisLeuPheIleProAlaLeuSerValGly 418
QY 1261 GGCATCACTGCGGAAGAAAGGGGACACATCAACAGCTGGGAGATTGGCGGAGCTCT 1320
Db 419 AlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSer 438
QY 1321 ATCAAGATTGCCCCCGGAGAGGCGCCAGACGTCAGCAAGAAAGATGTATCATCACTCCGCG 1380
Db 439 IleLysIleAlaProAlaGlnAlaProAspAlaLysValArgMetValIleIleThrGly 458
QY 1381 CCACCGAAGCCCACTTCAAGGCCAGGACCGATCTTTGGGAACTGAAAGAGGAAAC 1440
Db 459 ProProGlnAlaGlnPheLysAlaGlnIleArgIleTyrglyLysIleLysGlnGln 478
QY 1441 TTCTTTAACCACCAAGAGAGTGAAGTGAAGGCGATATCAAGTCCCTCTTCCACA 1500
Db 479 PheValSerProLysGlnGlnValLysLeuGlnAlaHisIleArgValProSerPheAla 498
QY 1501 GCTGGCCCGGATGATGGCAAGAGTGGCAAGACCGTGAAGCAACTGCAAGACTTAACA 1560
Db 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGlnLeuGlnHisLeuSer 518
QY 1561 GCGAGAGTCATCGTGCCTGTGACCAACCGCCAGATGAAATGAGAAAGTATCTGTGCA 1620
Db 519 AlaGlnValValAlaProArgAspGlnThrProAspGlnAsnAspGlnValValLys 538
QY 1621 ATTATCGGGCACTTCTTGTAGCCAGACTGACAGCCCAAGCCCAAGATAGAGAAATTGTACA 1680
Db 539 IleThrGlyHisPheTyraIaCysGlnValAlaGlnArgLysIleGlnGlnIleLeuThr 558
QY 1681 CAGGTGAAGCAGCAGCAGAGAA 1704
Db 559 GlnValLysGlnHisGlnGlnGln 566

```

```

US-09-897-778-348
; Sequence 348, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnetakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darlick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
;
Alignment Scores:
Pred. No.: 6,32e-143 Length: 579
Score: 1875.00 Matches: 369
Percent Similarity: 78.70% Conservative: 78
Best Local Similarity: 64.96% Mismatches: 95
Query Match: 31.86% Indels: 26
Gaps: 6
DB: 9

US-09-270-437D-8 (1-3283) x US-09-897-778-348 (1-579)
QY 73 ATGAAACAAGCTTTATCATGCGGAGAACCTGACCCCGCTCACCGCCAGACCTCCGCGAG 132
Db 1 MetAsnLysLeuTyrglyIleGlyAsnLeuSerGlnAsnAlaIleProSerAspLeuSer 20
QY 133 CTCCTTTGGGAGCAGAGAACTGCTGCGGAGAGAGCTGCTGCTGAAGTCCGGCTACGC 192
Db 21 IlePheLysAspAlaLysIleProValSerIleProPheLeuValLysThrGlyTyraIa 40
QY 193 TTCGAGACTACCCCGACAGACACTGCGCATCCGCGCATGAGACCTCTCCGGGTAA 252
Db 41 PheValAspCysProAspGlnSerIleProValSerIleProPheLeuValLysThrGlyTyraIa 60
QY 253 GTGGAATTGCACTGAGAAATATGAGAACTGATTACTCACTGCTTAAAGCTAAGAGAC 312
Db 61 IleGlnLeuHisGlyLysProIleGlnValGlnHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTTCAAGATTTCGAAATATCTCTCTCACTGCTGAGAGCTTTGATGAGACTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnIleProValLysAspSerLeu 100
QY 373 TTGGCTCAATATGGGACAGCTGAGAGATGGAACAAGTCAACACAGACAGACAGAAACCGCC 432
Db 101 LeuValGlnTyrglyValValGlnSerCysGlnGlnValAsnThrAspSerGlnThrAla 120
QY 433 GTTGTCAACGTCAATATGACAAACAGAGAAAGCAAAATAGCCATGAGAAAGCTAAGC 492
Db 121 ValValAsnValThrTyrsSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAGACTACTCTTCAAGATTTCCTTCAATCCCGGATGAAAGAGTGAAC 552
Db 141 GlyPheGlnLeuGlnAsnPheThrLeuLysValAlaTyrlleProAspGlnThrAlaAla 160
QY 553 TCCCCTTGGCCCTCTCAG-----CGAGCCAGAGT-----GGGACCACTCTTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgIleArgGlyLeuGlyGlnArgIleSer 180
QY 601 GAGCAAGGCCACGCCCTGTGGGGGCACTTCTCAAGGCCAGACAGATTGATTCCTCCGCTGGCG 660

```


Db	181	AngInGly----	SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199
QY	661	ATCTGGATCCCA	CCAGTTTGTGGTGCATCATCGGAAAGAGGGCTTGACATTAAG	720
Db	200	LeuLeuValPro	ThrGlnPheValGlyAlaIleIleGlyLysGlnGlyAlaThrIleArg	219
QY	721	AAACATCACTA	AGACCCAGTCCCGGATGATATCCATGAAAGAAAGAACTCTGAGCT	780
Db	220	AsnIleThrLys	GlnThrGlnSerLysPheValHisArgLysGlnAsnAlaGlyAla	239
QY	781	GGAGAGAACCT	GTACCATTCATGCCACCCAGAGGGGACTTCTGAAAGCATGCCGATG	840
Db	240	AlaGlnLysSer	IleThrIleLeuSerThrProGlnGlyThrSerAlaAlaCysLysSer	259
QY	841	ATTCCTGAAT	CATCATGACAGAGGACAGTATGACCAAACTGACGGAGAGATTCTCTG	900
Db	260	IleLeuGlnIle	MetHisLysGlnAlaGlnAspIleLysPheThrGlnGlnIleProLeu	279
QY	901	AAATCTTGGCA	CACATGCTTGTTGGTGGAAACACTGATTTGAAAGAGGACGAATTTG	960
Db	280	LysIleLeuAla	HisAsnAsnPheValGlyArgLeuIleGlyLysGlnGlyArgAsnLeu	299
QY	961	MAGAAATTA	GAACTGAAACAGGGACCAAGATACATCTCATCTTTCAGAGATTGAGC	1020
Db	300	LysLysIleGln	GlnAspThrAspThrLysIleThrIleSerProLeuGlnGlnLeuThr	319
QY	1021	ATATCAACCC	CGGAAAGAACCTCATCTGTGAAGGGCACAGTTGAGCGCTGTGCCAGTCT	1080
Db	320	LeuTyraenPro	GlnArgThrIleThrValLysGlyAsnValGlnThrCysAlaLysAla	339
QY	1081	GAGATAGAAT	TATGAAAGAGCTGCGTGAAGGCTTTGAAATATGATCTGGCTGTATAC	1140
Db	340	GlnGlnGlnIle	MetLysLysIleArgGlnSerTyraenAspIleAlaSerMetAsn	359
QY	1141	ACCCACTCC--	-----GGATCTTC-----	1158
Db	360	LeuGlnAlaHis	LeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProProThr	379
QY	1159	-----	-----TCCAGCCTGTACCCCCATACCCAGTTTGCCCGCTTCCGCAT	1200
Db	380	SerGlyMetPro	ProProThrSerGlyProProSerAlaMetThrProProTyProGln	399
QY	1201	CATCACTCTTA	TCGACAGACAGAGATTTGATCTCTTCATCCCAACCCAGGCTGTGGGC	1260
Db	400	PheGlnGln---	SerGlnThrGlnThrValHisIlePheIleProAlaLeuSerValGly	418
QY	1261	GCCATCATCG	GGAGAGAGGGGACACATCAAACAGCTGGCGAGATTGCCGAGCCTT	1320
Db	419	AlaIleIleGly	LysGlnGlyGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSer	438
QY	1321	ATCAAGATTG	CGCCCGGAGGCGCCAGACGTCAAGCGAAAGAGTGTATATATACCGGG	1380
Db	439	IleLysIleAla	ProAlaGlnAlaProAspAlaLysValaLysMetValIleIleThrGly	458
QY	1381	CCACCGAAGC	CAAGTTCAGAGCCACAGGACGGAATCTTGGGAAACGTGAAGAGAAAC	1440
Db	459	ProProGlnAla	GlnPheLysAlaGlnGlyArgIleTyraenLysIleLysGlnGlnAsn	478
QY	1441	TTCTTTAAAC	CCCAAGAAAGTAGAGCTGGAAGCGCATATCAAGTGCCTCTTCACA	1500
Db	479	PheValSerPro	LysGlnGlnValLysLeuGlnAlaHisIleArgValProSerPheAla	498
QY	1501	GCTGGCGGGG	ATGTTGGCAAGGATGGCAAGCCGTGAACGAATCTGCGAATTAACAGT	1560
Db	499	AlaGlyArgVal	IleGlyLysGlyLysThrValAsnGlnLeuGlnAsnLeuSerSer	518
QY	1561	GCAGAAGTCAT	CGTGCCTCTGTACCAACCGCATATAAATGAGAAAGTATCGTGCAGA	1620
Db	519	AlaGlnValVal	AlaProArgAspGlnThrProAspGlnAsnAspGlnValValVal	538
QY	1621	ATTATCGGGCA	CTTCTTTGCTAGCCAGACTGCACAGCGGCAAGATCGAGAAATTTGACAA	1680

[illegible]

```

Db      141 GlyPheGlnLeuGlnAsnPheThrLeuIysValAlaIleProAspGluThrAlaIa 160
QY      553 TCCCTTCGCCCCCTCAG-----CGAGCCAGCGT-----GGGAGCACTCTTCCCGG 600
Db      161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySer 180
QY      601 GAGCAGAGCCACGCCCCCTGGGGGCACTTCTCAGGCCACAGATGATTCCCGCTGGCG 660
Db      181 ArgGlnGly---SerProGlySerValSerIysGlnIysProCysAspLeuProLeuArg 199
QY      661 ATCTGTGTCGCCACCCAGTTGTGTGTGGTCATCGAAGAGAGGGCTTGACATTAAG 720
Db      200 LeuLeuValProThrGlnPheValGlyIleIleGlyIysGlnGlyAlaIleThrIleArg 219
QY      721 AACATCATAGCAGACCCAGTCCCGGGTAGATTATCCATGAAAAAGAACTCTGGAGCT 780
Db      220 AsnIleThrIysGlnThrGlnSerIysIleAspValHisArgIysGlnAsnAlaGlyAla 239
QY      781 GCAGAGAAAGCCGTGCACCATCATGTCACCCGAGGGGACTTGAAGCATCGCCGATG 840
Db      240 AlaGlnIysSerIleThrIleLeuSerThrProGlnGlyThrSerAlaAlaCysIysSer 259
QY      841 ATTTCTTGAATCATGCAGAAAGAGGACAGATGAGACCAACTAGCCGAGAGATTCTCTG 900
Db      260 IleLeuGlnIleMetHisIleIysGlnAlaGlnAspIleIysPheThrGlnGlnIleProLeu 279
QY      901 AAAATCTTGGCACCAATGGCTTGGTGAAGAATGATTGAAAAAGAGGCGAAATTG 960
Db      280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlnIysIleArgAsnLeu 299
QY      961 AAGAAATGTAACATGAAACAGGACGACCAAGATTAACATCTCTTGTGAGATTGAGC 1020
Db      300 LysIysIleGlnGlnAspThrAspThrIysIleThrIleSerProLeuGlnGlnLeuThr 319
QY      1021 ATATCAACCCCGAAAGAACCATCATCTGTGAAGGCGACAGTTGAGCGCTTGCCAGTGT 1080
Db      320 LeuIysAsnProGlnArgThrIleThrValIysGlyAsnValGlnThrCysAlaIysAla 339
QY      1081 GAGATGAGATTTATGAAGAAGCTGGCTGAGGCTTGAAGAATGATGCTGGCTGTTAC 1140
Db      340 GlnGlnGlnIleMetLysIleArgGlnSerIysGlnAsnAspIleAlaSerMetAsn 359
QY      1141 ACCCACTCC-----GGATACCTTC----- 1158
Db      360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPheProProThr 379
QY      1159 -----TCCAGCCTGTACCCCGCATCACCAAGTTGGCCGCTTCCCGCAT 1200
Db      380 SerGlyMetProProProThrSerGlyProProSerAlaMetThrProProIysProGln 399
QY      1201 CATCACTCTTATCCAGAGAGATTTGAAATCTCTCATCCCAACCCAGGCTGGGGC 1260
Db      400 PheGlnGln---SerIleThrGlnThrValHisLeuPheIleProAlaLeuSerValGly 418
QY      1261 GCCATCATCGGAGAAAGAGGGGACACATCAACAGCTGGCGAGATTGCGCGAGCCTCT 1320
Db      419 AlaIleIleGlyIysGlnGlnIleGlnHisIleIysGlnLeuSerThrPheAlaIleAsp 438
QY      1321 ATCAAGATTGGCCCTCGGAAAGGCCAGACGTGAGGAAAGATGGTCAATCATCCGGG 1380
Db      439 IleIysIleAlaProAlaGlnAlaProAspAlaIysValArgMetValIleIleIleArg 458
QY      1381 CCACCGGAAGCCAGTTTCMAAGGCCCGAGGACGAGATCTTTGGGAACTGAAAGAGAAAC 1440
Db      459 ProProGlnAlaIlePheIysAlaGlnGlyArgIleIysGlyIysIleIysGlnGlnAsn 478
QY      1441 TTCTTAATCCCAAGAAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTTCCACA 1500
Db      479 PheValSerProIysGlnGlnValIysIleGlnAlaHisIleIleArgValProSerPheAla 498
QY      1501 GCTGGCCGGGTGATTGGCAAGGTGGCAGACCGTGAACGAACTGCAGAACTTAACAGT 1560
Db      499 AlaGlyArgValIleGlyIysGlyIysThrValAsnGlnLeuGlnAsnLeuSer 518

```

```

QY      1561 GCAGAGATCGTGCCTCGTACCAAGCCAGATGAATAATGAGAGATGATCTCAGA 1620
Db      519 AlaGlnValValProArgAspGlnThrProAspGlnAsnAspGlnValValIys 538
QY      1621 ATTAATCGGGCACTTCTTTGCTAGCCAGACTGCACGCCCAAGATCAGGAAATGTACAA 1680
Db      539 IleThrGlnHisPheIysAlaCysGlnValAlaGlnArgIysIleGlnGlnIleLeuThr 558
QY      1681 CAGGTGAAGCCAGCAGAGCAAGAA 1704
Db      559 GlnValIysGlnHisGlnGlnGln 566

RESULT 13
US-09-897-778-449
/ Sequence 449: Application US/09897778
/ Patent No. US20020147143A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tonglong
/ APPLICANT: Marnierakis, Margarita
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Veddyck, Thomas S.
/ APPLICANT: Carter, Darrick
/ APPLICANT: Watanabe, Yoshihiro
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: Peckham, David W.
/ APPLICANT: Fanger, Neil
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.455C16
/ CURRENT APPLICATION NUMBER: US/09/897,778
/ CURRENT FILING DATE: 2001-06-28
/ NUMBER OF SEQ ID NOS: 467
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 449
/ LENGTH: 579
/ TYPE: PRU
/ ORGANISM: Homo sapiens
US-09-897-778-449

Alignment Scores:
Pred. No.: 6,32e-143 Length: 579
Score: 1875.00 Matches: 369
Percent Similarity: 78.70% Conservative: 78
Best Local Similarity: 64.96% Mismatches: 95
Query Match: 31.86% Indels: 26
DB: 9 Gaps: 6

US-09-270-437D-8 (1-3283) x US-09-897-778-449 (1-579)
QY      73 ATGAACAAGCTTTATCATCGGAACTGAGCCCGCGCTCACCGCGAGCACTCCGAGCAG 132
Db      1 MetAsnLysLeuIleIleGlyAsnLeuSerGlnAsnAlaAlaProSerAspLeuGlnSer 20
QY      133 CTCTTTGGGACAGAGAAGCTGCGCCGCGGACAGAGTCTGTGTAAGTCCGCTAGGCC 192
Db      21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValIysThrGlyTyrAla 40
QY      193 TTCTGGAGCTACCCCGACCAAGAACTGGGCGCATCGGCGCATCGAGACCCCTCGGCTAA 252
Db      41 PheValAspCysProAspIleSerThrAlaLeuLysAlaIleGlnAlaLeuSerGlyIys 60
QY      253 GTGGAATTGCATGGGAAATTCATGAAAGTGTACTAGTCTCTAATAAGCTAAGAGC 312
Db      61 IleGlnLeuHisIleLysProIleGlnValGlnHisSerValProIysArgGlnArgIle 80
QY      313 AGGAAATTCAGATTGGAATCATCCCTGCTCACTGAGTGGAGGAGGTGTGATGACCT 372
Db      81 ArgIysLeuGlnIleArgAsnIleProIleProHisIleGlnIleProIysArgGlnArgIle 100
QY      373 GTGCTCATATGAGGACAGTGAAGATGTGGAACAAGTGAACAAGTGAACAAGTGAACA 432
Db      101 LeuValGlnIysIysValValGlnIysSerCysGlnGlnValAsnThrAspSerGlnThrAla 120

```

```

QY 433 GTTTCACGTCACATATGCAACAGAGAGAGCAAAAATAGCCATGAGAGCTAAC 492
DB |||||.....:|||||
QY 121 ValIleAsnValThrTyrSerSerIysAspGlnAlaArgGlnAlaLeuAspIysLeuAsn 140
DB |||||.....:|||||
QY 493 GGGCATCACTTGGAGAACTACTCTTCAAGATTCTCCATCCTCCGATGAAGAGTGAAC 552
DB |||||.....:|||||
QY 141 GlyPheGlnLeuGlnAsnPheThrLeuIysValAlaTyrIleProAspIleThrAlaIa 160
DB |||||.....:|||||
QY 553 TCCCTTGGCCCCCTCAG-----CGAGCCAGCGT-----GGGACCACTCTTCCCG 600
DB |||||.....:|||||
QY 161 GlnGlnAsnProLeuGlnIleProArgGlyArgArgGlyLeuGlyGlnArgIleArgSer 180
DB |||||.....:|||||
QY 601 GAGCAGAGCCAGCCCTGCGGCACTTCTCAGCCAGCAGACAGATTGATTCCCGCTGG 660
DB |||||.....:|||||
QY 181 ArgGlnIly---SerProGlySerIleSerIysGlnIysProCysAspLeuProLeuArg 199
DB |||||.....:|||||
QY 661 ATCTCGTGTCCCAACCCAGCTTTGTGTGTCATCATCGAGAAAGAGGCGCTTGAACATAA 720
DB |||||.....:|||||
QY 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyIysGlnGlyAlaThrIleArg 219
DB |||||.....:|||||
QY 721 AACATCACTAAGCAGAGCCAGCTCCGGGTAGATATCCATAGAAAAGAACTCTGGAGCT 780
DB |||||.....:|||||
QY 220 AsnIleThrIysGlnThrGlnSerIysIleAspValHisArgIysGlnAsnAlaGlyAla 239
DB |||||.....:|||||
QY 781 GCAGAGAGACCTGTACCATCCATCCACCCAGAGGAGGACTTCTGAGACATGCGCGCATG 840
DB |||||.....:|||||
QY 240 AlaGlnIysSerIleThrIleLeuSerThrProGlnGlyThrSerAlaIleCysIysSer 259
DB |||||.....:|||||
QY 841 ATTCTTGAATCATGCGAGAAAGAGGCGAGATAGACCAATAGCCGAGAGATTCTCTG 900
DB |||||.....:|||||
QY 260 IleLeuGlnIleMetHisIysGlnAlaGlnAspIleIysPheThrGlnIleProLeu 279
DB |||||.....:|||||
QY 901 AAAATCTTGGCAGCAATGCTTGTGGAGAGCTGATGTGAAAAGAGAGCAAAATTG 960
DB |||||.....:|||||
QY 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyIysGlnGlyArgAsnLeu 299
DB |||||.....:|||||
QY 961 AAGAAATTTGAATCACTGAACAGGAGCAAGATAACAATCTCTTGGACAGATTGAGC 1020
DB |||||.....:|||||
QY 300 LysIysIleGlnIleAspThrIleThrIleThrIleSerProLeuGlnIleLeuThr 319
DB |||||.....:|||||
QY 1021 ATATACAAACCCGGAGAAAGACATATCACTGTGAAGGCGACAGTGAAGCCTGTGCCAGTCT 1080
DB |||||.....:|||||
QY 320 LeuIysAsnProGlnIleArgThrIleThrValIysGlnAsnValGlnThrCysAlaIysAla 339
DB |||||.....:|||||
QY 1081 GAGATAGAGATTATGAGAGAGCTGCGAGCCCTTGAATAATGATATGCTGGCTGTAAAC 1140
DB |||||.....:|||||
QY 340 GlnIleGlnIleMetIysIleArgGlnSerTyrGlnAsnAspIleAlaSerMetAsn 359
DB |||||.....:|||||
QY 1141 ACCCACTGC-----GGATATCTTC----- 1158
DB |||||.....:|||||
QY 360 LeuGlnAlaHisIleLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProThr 379
DB |||||.....:|||||
QY 1159 -----TCCAGCTGTACCCCATCAGCAAGTTGGCCGTTCCCGCAT 1200
DB |||||.....:|||||
QY 380 SerGlyMetProProProThrProThrSerGlyProProSerAlaMetThrProProTyrProGln 399
DB |||||.....:|||||
QY 1201 CATCATCTTATCCAGAGAGGAGATGTGATCTTCAATCTCCACCCAGCGCTGTGGC 1260
DB |||||.....:|||||
QY 400 PheGlnIle---SerGlnThrGlnThrValHisIleuPheIleProAlaLeuSerValGly 418
DB |||||.....:|||||
QY 1261 GCCATCATCGGAGAGAGGCGGACACATCAAAACAGCTGCGAGATTGCGCGAGCCTCT 1320
DB |||||.....:|||||
QY 419 AlaIleIleIleGlyIysGlnGlyIleHisIleIysGlnLeuSerArgPheAlaGlyAlaSer 438
DB |||||.....:|||||
QY 1321 ATCAAGATTGCCCCCTGCGAGAGGCCCAAGCTCAGGAGAAAGAGTGCATCATCACCGG 1380
DB |||||.....:|||||
QY 439 IleIysIleIleAlaProAlaGlnAlaProAspAlaIysValArgMetValIleIleThrGly 458
DB |||||.....:|||||
QY 1381 CCACCGAGAGCCAGTTCAAGCGCCAGGAGCGATCTTTGGAGAACTGAAAGAGAGAAAC 1440
DB |||||.....:|||||
QY 459 ProProGlnAlaGlnPheIysAlaGlnGlyArgIleTyrIleTyrIysIleIysGlnIleAsn 478

```

```

QY 1441 TTCTTAAACCCCAAGAGAGAGTGAAGCTGAAGGCGCATATCAAGTGCCTCTTCCACA 1500
DB |||||.....:|||||
QY 479 PheValSerProIysGlnGlnValIysLeuGlnAlaHisIleArgValProSerPheAla 498
DB |||||.....:|||||
QY 1501 GCTGGCCCGGATGATTTGGAGAAAGTGGCAAGACCGGAGAGCAAACTGCAGAACTTAACCG 1560
DB |||||.....:|||||
QY 499 AlaGlyArgValIleGlyIysGlyIysGlyIysThrValAsnGlnLeuGlnAsnLeuSer 518
DB |||||.....:|||||
QY 1561 GCAGAAATCATCTGTGCTCGTACCAACGCGAGATGAATAATGAGAAAGTGAATGTCAGAA 1620
DB |||||.....:|||||
QY 519 AlaGlnValIleValProArgAspIleThrProAspGlnAsnAspGlnValValIys 538
DB |||||.....:|||||
QY 1621 ATTATCGGCGCACTTTTGTGCTAGCCAGACTGCACGCGCCAGATCAGGAAATTTGATCAA 1680
DB |||||.....:|||||
QY 539 IleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgIysIleGlnGlnIleLeuThr 558
DB |||||.....:|||||
QY 1681 CAGGTGAAGCAGACGAGCAGAGAA 1704
DB |||||.....:|||||
QY 559 GlnValIysGlnHisGlnGlnGln 566
DB |||||.....:|||||

RESULT 14
US-10-007-700-348
; Sequence 348, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margalita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-348

Alignment Scores:
Pred. No.: 6.32e-143
Score: 1875.00
Percent Similarity: 78.70%
Best Local Similarity: 64.96%
Query Match: 31.86%
DB: 12
Gaps: 6

US-09-270-437D-8 (1-3283) x US-10-007-700-348 (1-579)

QY 73 ATGAGCAAGCTTACATCGGAGACCTGAGGCCGCCGTACCGCCGACGACCTTCGGCAG 132
DB |||||.....:|||||
QY 1 MetAsnIysLeuIyrIleGlyAsnLeuSerGlnAsnAlaIleProSerIlePheGlnIle 20
DB |||||.....:|||||
QY 133 CTCCTTTGGGAGACAGAGAGCTGCCCCCTGCGGAGCAGAGTCTGTGTAAGTCCGCTACGCC 192
DB |||||.....:|||||
QY 21 IlePheIysAspAlaIysIleProValSerGlyProPheLeuValIysThrGlyTyrAla 40

```

```

QY 193 TTGGTGAACACCCGACAGAACTGGGCGCATCCGCGCATCGAACCCCTCTGGGGTAA 252
DB 41 PheValaspCysProaspGlnSerThrValLeuValAlaIleGlnAlaLeuSerGlyLys 60
QY 253 GTGGAATTCATGGGAAAATATGGAAGTTGATTATTCAGTCTCTAAATAAGCATGAGAC 312
DB 61 IleGlnLeuHisGlyLysProIleGlnValGlnHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTGGAACATCCCTCTCACTGCACTGGGAGGCTGTGGATGACTT 372
DB 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnThrGlnValLeuAspSerLeu 100
QY 373 TTGGCTCAATTCGGACACGTGGAAGATGGAACAAGTCAACACAGACAGAAACCGCC 432
DB 101 LeuValGlnTyrGlyValValGlnSerCysGlnGlnValAsnThrAspSerGlnThrAla 120
QY 433 GTTGTCAACGTCACTATGCAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
DB 121 ValValAlaSerValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTGAGAACTACTCTTCAAGATTCTCTACATCCCGATGGAAGAGTGAAC 552
DB 141 GlyPheGlnLeuGlnAsnPheThrLeuLysValAlaTyrIleProAspGlnThrAlaAla 160
QY 553 TCCCTTGGCCCTCCAG-----CGAGCCAGCGT-----GGGAGACCATCTTCCCGG 600
DB 161 GlnGlnAsnProLeuGlnProArgLysArgGlyValGlnArgGlySerSer 180
QY 601 GAGCAAGGCCACGCGCTGGGGGCACTTCTCAAGCCAGACAGATTGATTTCCCGTCCG 660
DB 181 ArgGlnGly---SerProLysSerValSerLysGlnLysProCysAspLeuArg 199
QY 661 ATCTGTGCTCCCAACCACTTTGTTGTGTCATCATCGAAAGAGGGCTTGACCAATAG 720
DB 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGlnGlyAlaThrIleArg 219
QY 721 AACATCACTAAGCAAGCCAGTCCGCGGATGATCCATCAGAAAGAACTCTGAGCT 780
DB 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGlnAsnAlaGlyAla 239
QY 781 GCAGAGAGACCTGTGCATCCATGCCACCCAGAGGGAGCTTCTGAGCATGCCGATG 840
DB 240 AlaGlnLysSerIleThrIleLeuSerThrProGlnGlyThrSerAlaIaCysLysSer 259
QY 841 ATCTTGAATCATGCAAGAAAGGCGATGAGACCAAACTAGCCCAAGAGATCTCTCG 900
DB 260 IleLeuGlnIleMetHisLysGlnAlaGlnAspIleLysPheThrGlnGlnIleProLeu 279
QY 901 AAAATCTGGCAACAATGGCTGTGGAAAGCTGATGGAAGAGAGAGAGAGAGAGAG 960
DB 280 LysIleLeuAlaHisAsnAsnProValGlyArgLeuIleGlyLysGlnGlyArgAsnLeu 299
QY 961 AAGAAATTCAGATGAAACAGAGGAGCAAGATTAACATCTCATCTTTCAGATTGAGC 1020
DB 300 LysLysIleLeuGlnAsnThrAspThrLysIleThrIleSerProLeuGlnGlnLeuThr 319
QY 1021 ATATACACCCGAAAGAACATCATCTGTGAAGGCACTTGAAGCGCTTGACAGTCT 1080
DB 320 LeuTyrAsnProGlnArgThrIleThrValLysGlyAsnValGlnThrCysAlaLysAla 339
QY 1081 GAGATGAGATTATGAAAGAGCTGCGAGAGCTTTGAAATGATATGCTGCTGTAAAC 1140
DB 340 GlnGlnIleThrMetLysLysIleArgLysSerTyrGlnAsnAspIleAsnMetAsn 359
QY 1141 ACCCACTCC-----GGATACCTC----- 1158
DB 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProThr 379
QY 1159 -----TCCAGCTGTACCCCATACACAGATTGGCCGCTTCCGCGCAT 1200
DB 380 SerGlyMetProProThrThrSerGlyProProSerAlaMetThrProProTyrProGln 399
QY 1201 CATCACTCTTATCCAGAGAGAGATTTGTAATCTTTCATCCCAACCCAGAGCTGTGGGC 1260

```

```

DB 400 PheGlnGln---SerGlnThrGlnThrValHisLeuPheIleProAlaLeuSerValGly 418
QY 1261 GCCATCATCCGGAAGAGAGGGGCAACATCAACAGCTGGGAGATTCCGCGAGCTCT 1320
DB 419 AlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSer 438
QY 1321 ATCAAGATTGCCCTCGCGAAGGCCACAGCTCAACGAGAGAGATGCTCATCATCCGGG 1380
DB 439 IleLysIleAlaProAlaGlnAlaIleProAspAlaLysValArgMetValIleIleThrGly 458
QY 1381 CCACCGGAGCCCACTTAAAGGCCCGAGAGAGATCTTGGGAACAGAAAGAGAAAC 1440
DB 459 ProProGlnAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLysGlnLysn 478
QY 1441 TTCTTAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTTTCACA 1500
DB 479 PheValSerProLysGlnGlnValLysLeuGlnAlaHisIleArgValProSerPheAla 498
QY 1501 GCTGCGCGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAATCGACAACTTAACAGT 1560
DB 499 AlaGlyArgValIleGlyLysGlyLysThrValAlaGlnGlnLeuGlnAsnLeuSerSer 518
QY 1561 GCAGAACTCATGCTGCTCTGACCAAGCCAGATGAAATGAGAGAGTATCGTCA 1620
DB 519 AlaGlnValAlaValProArgAspGlnThrProAspGlnAsnAspGlnValValLys 538
QY 1621 ATTATCGGGCACTTTGCTTGGCCAGACTGCACAGCCGCAAGATCAGGAAATTGTACA 1680
DB 539 IleThrGlyHisPheThrAlaCysGlnValAlaGlnAlaArgLysIleGlnGlnIleLeuThr 558
QY 1681 CAGGTGAAGCAGAGAGAGAGAA 1704
DB 559 GlnValLysGlnHisGlnGlnGln 566

```

```

RESULT 15
US-10-007-700-446
; Sequence 446, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yaeli A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Ranger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Derrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Beckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Fey, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 446
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-446

```

```

Alignment Scores: 632e-143 Length: 579
Pred. No.: 1875.00 Matches: 369
Score:

```

Percent Similarity: 78.70% Conservative: 78
 Best Local Similarity: 64.96% Mismatches: 95
 Query Match: 31.86% Indels: 26
 DB: 12 Gaps: 6

US-09-270-437d-8 (1-3283) x US-10-007-700-446 (1-579)

QY 73 ATGAAACACCTTTACATCGGAAACCTGAGCCCGCGTACCGCGACGACCTCCGGCAG 132
 |||
 Db 1 MetAsnLysLeuTrpIleGlyAsnLeuSerGluAsnAlaIleProSerSerLeuSer 20

QY 133 CTCCTTTGGGAGACGAGAGCTGCGCCCTGCGGAGACAGTCTGCTGTAAGTCCGCTACGCC 192
 |||
 Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40

QY 193 TTCTGGGACTACCCCGACGACCAAGTGGCCATCCCGCCATGAGAACCTCTCGGATAA 252
 |||
 Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60

QY 253 GTGAAATTGCAATGGGAAATCATGGAAGTGAATTGATCTGCTGTAAGTCTGTAAGAAC 312
 |||
 Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGluArgIle 80

QY 313 AGGAAATTCAGATTGGAACATCCCTCTCCTGACCTGAGTGGAGAGTGTGATGAGACTT 372
 |||
 Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100

QY 373 TTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAGTCAACACAGACAGAAACCGCC 432
 |||
 Db 101 LeuValIleGlnTyrGlyValIleValIleSerCysGluGlnValAsnThrAspSerGluThrAla 120

QY 433 GTTGTCACGTCACATATGCAACAGAGAAAGAGAAATAGCCATGAGAAAGTAAAGC 492
 |||
 Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140

QY 493 GGGCATCAGTTTGAAGACTACTCTTCAAGATTCTTCAATCCCGATGAGAGAGGTAGC 552
 |||
 Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160

QY 553 TCCCTTGGCCCTCAG-----CGAGCCCAAGCGT-----GGGACCACTCTTCCCGG 600
 |||
 Db 161 GlnIleAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180

QY 601 GAGGAGGCGACGCGCCCTGCGGAGCACTTCTCAAGCGCAGACAGATGATTTCCCGCTGCG 660
 |||
 Db 181 ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199

QY 661 ATCTGTGTCGCCACCCAGTTTGTGTGTCATCATCGGAAAGAGGGCTTGACCATTAAG 720
 |||
 Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219

QY 721 AACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGAGCT 780
 |||
 Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239

QY 781 GCAGAGAGGCTGTCACATCCATGCAACCCGACAGGGGAGATTCTGAAAGCATGCCCATG 840
 |||
 Db 240 AlaIleLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259

QY 841 ATTCTTGAATTCATGACAGAAAGAGAGAGATGACCAAACTAGCCGAGAGATTCTCTG 900
 |||
 Db 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279

QY 901 AAAATCTTGGCAGCAATGGCTTGGTGAAGACTGATGAAAAAGAGCAGAAATTTG 960
 |||
 Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299

QY 961 AAGAAATTTGAACATGAAACAGGAGCAAGATTAATCATCTTCTGACAGATTTGAGC 1020
 |||
 Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319

QY 1021 ATATACAAACCCGAAAGAACCATCATGTGAAGGGCAGATTGAGGCTGTGACAGTGTCT 1080
 |||

Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339

QY 1081 GAGATAGAGATTATGAGAGAGCTGCGAGCCCTTGAAGAAATGATATGCTGCTGTTAAC 1140
 |||
 Db 340 GlnGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359

QY 1141 ACCCACTCC-----GGAATCTTC----- 1158
 |||

Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProThr 379

QY 1159 -----TCCAGCCTGTAACCCCATCACAGATTGGCCCTTCCCGCAT 1200
 |||

Db 380 SerGlyMetProProProThrSerGlyProProSerAlaMetThrProProTyrProGln 399

QY 1201 CATCACTCTTATCCAGACGAGAGATTGTGAATCTTTCATCCCAACCCAGGCTGCGGCG 1260
 |||

Db 400 PheGluGln---SerGluThrGluThrValHisPhePheIleProAlaLeuSerValGly 418

QY 1261 GCCATCATCGGAGAGAGAGGCGGACACATCAAAACAGCTGGCGAGATTGCGCGAGCCTCT 1320
 |||

Db 419 AlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAlaGlyAlaSer 438

QY 1321 ATCAAGATTGCCCCCGGAGAGGCCACAGACCTGACCGAAAGATGTCTATCATCCCGG 1380
 |||

Db 439 IleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIleIleThrGly 458

QY 1381 CCACCGAAGCCAGTTCAAGGCCGAGAGGAGATCTTGGGAAACTGAAAGAGAAAC 1440
 |||

Db 459 ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLysGluLysAsn 478

QY 1441 TTCTTTAAACCCCAAGAAAGAGTGAAGCGCATATCAGAGTCCCTCTTCCACA 1500
 |||

Db 479 PheValSerProLysGluGluValLysLeuGlnAlaHisIleArgValProSerPheAla 498

QY 1501 GCTGCGCGGCGATTTGCAAGAGTGGCAAGACCGTGAACCAAGTCCGACAACTTAACAGT 1560
 |||

Db 499 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuSerSer 518

QY 1561 GCAGAGATCATCGTGCCTCGTGAACCAACGCGCAGATGAAGAAATGAGGAAGTATGTCAGA 1620
 |||

Db 519 AlaGluValValIleProArgAspGlnThrProAspGluAsnAspGluValValLys 538

QY 1621 ATTATCGGCGCATCTTCTTGTAGCCAGACTGACACAGCGCAAGATCAGGAAATGTACAA 1680
 |||

Db 539 IleThrGlyHisPheTyrAlaCysGluValAlaGlnArgLysIleGlnGluIleLeuThr 558

QY 1681 CAGGTGAAGCAGCAGAGCAAGAAA 1704
 |||

Db 559 GlnValLysGlnHisGlnGln 566

Search completed: July 23, 2004, 11:58:03
 Job time : 192.647 secs


```

QY 520 AAGATTCTTCATCCCGGATGAAAGGTGAGCTCCCTTGCCGCCCTCAGCGAGCCAG 579
Db 283 -----Gln 283
QY 580 CGTGGGAGCACTCTCCCGGAGCAAGGC-----CAGGCCCTGAGGAGCACTTCTCAG 633
Db 284 HisArgAsnHisAsnGlnSerHisSerGlyProHisHisIleProGlnAsnLeuMet 303
QY 634 GCCAGACAGATT-----GATTCCCGCTGGGATCTCTGCTCCACCCAGCTTCTGTGT 687
Db 304 ProArgCysMetLeuIleAspTyrProIleArgCysValValGlnGlyLysThrHisAla 323
QY 688 GGCATCATCGGAAAGAGAGGCTTGCATTAAGACATCACTCAACACACACAGCCCGCG 747
Db 324 ValIleIleGlyProAsnGlySerThrIleLysAspIleHisSerSerThrArgCysArg 343
QY 748 GTAGAT-----ATCCATAGAAAGAAAGAACTCT-----GAGCTGCAGAGAGACCT 792
Db 344 ValAspPheValAsnLeuSerLysGlnArgThrValIleGlnIleAsnAsnAspArgIle 363
QY 793 GTCACCATCATCCAGCCAGAGGAGACTTGTAGAGCATGCCGAGATGATTCTTGAATC 852
Db 364 LeuThrValHisGlyValAlaGlnGlnAlaThrLysAlaValAlaArgIleLeuAspVal 383
QY 853 ATGCAGAAAGAGGCA-----GATGAGACCAAACTAGCCGAGAGATTCTCTGAATC 906
Db 384 IleGlnSerGlnValValLysAspPheValAsnValGlyAlaAspThrValLeuArgMet 403
QY 907 TTGGCACAACATGCTGTGGTGAAGACTGATTGAAAGAAAGAGCAAAATTTGAAGAA 966
Db 404 ArgAlaHisAsnGlnLeuGlySerGlyArgLeuIleGlyLysAlaGlySerSerLysGln 423
QY 967 ATTGAACATGAAGAGGAGCAAGATTAAGATCTGTCAGAGATTGAGCATATAC 1026
Db 424 IleMetGlnLysThrGlyThrAsnIleThrValThrLysTyr-----IleGln 439
QY 1027 AACCCG-----GAAAGAAC 1041
Db 440 ProProGlyGlyIleSerGlyLeuThrAlaAsnGlnLeuGlyLeuMetGlnArgThr 459
QY 1042 ATCACTGTGAAGGC-----ACAGTTGAGCCTGTGCAGTCTGAGATAGATTAAGAAG 1098
Db 460 IleMetValArgGlyProSerIleGlnValAlaGlnAlaGlnAlaLeuIleSerAla 479
QY 1099 AAGCTCGCGAGGCGCTTGAATATGAT----- 1125
Db 480 LysLeuLysLysCysTyrGlnSerAspSerGlnLeuArgAlaGlnSerMetGlnCysPro 499
QY 1126 -----ATGCTGCTGTTTACACCCACTCCGATATTCTTCAGCCTGTATCCCC 1173
Db 500 MetProProMetMetMetProProIleLeuProProGlyAlaSerSerSerAlaValSer 519
QY 1174 CATCACCAGTTGGCCCGTTCG-----CATATCACTTTTATCCAGAG 1218
Db 520 AlaProHisPheIleProThrProValGlyValMetGlnValGlnHisPheAlaSerSer 539
QY 1219 CAGGAGATTGTGAAT----- 1233
Db 540 GlnHisLeuValHisGlnAsnAlaAsnAsnSerPheLeuGlnProGlyValLeuGlnIle 559
QY 1234 -----CTCTTCATCCCAACCCAGCGCTGTGGGC 1260
Db 560 GlnProGlyThrThrAsnLeuArgGlnValArgMetTyrAlaProAspSerMetIleGly 579
QY 1261 GCCATCATCGGAAAGAGGCGGACACATCAACACAGTGGCGAGTTCCGCCGAGCTCT 1320
Db 580 AlaLeuIleGlyAlaLysGlyLysAsnIleLysMetIleLeuArgAspThrGlyAlaSer 599
QY 1321 ATCAAGATT-----GCCCTCGGAGAGGCCAGAC----- 1350
Db 600 ValLysIleGlnAlaProGlnGlnLysThrGlnArgGlnAlaGlnGlnAlaGlnLysLys 619
QY 1350 ----- 1350

```

```

Db 620 ArgLysLeuAspGlnThrAspSerGlyCysGlnGlyValAlaAsnGlyAspHisProGln 639
QY 1350 ----- 1350
Db 640 GlnPheLeuGlnAspAsnAlaThrIleAsnSerSerAspAlaIleGlnGlnLysProLys 659
QY 1351 -----GTACAGCAAGAGATGATCATCATCACCGGCCACCGGAGGCCAGTCAAGGCCAG 1407
Db 660 ProValSerGlnArgMetValThrIleAsnGlnAspPheLeuGlnLeuLysAlaGln 679
QY 1408 GAGCGATCTTTGGGAAACTGAAAGAGGAAACCTTTAAACCC----- 1452
Db 680 SerTyrValPheSerLysIleAlaGlnThrSerSerSerLeuProSerSerGlyMetAsp 699
QY 1453 -----AAGAGAAAGATGAGCTGGAAGCGCATATCAGATGCCCTCTTCCACAGCT 1503
Db 700 GlyAspArgSerHisMetLeuArgIleArgThrGlnValSerValProThrArgIleIle 719
QY 1504 GCGCCGGTGAATTTGGCAAGAGTGGCAAGACCGTGAACGAACTGCAGAACTTAACAGTGA 1563
Db 720 GlyArgIleIleGlyLysGlyGlyGlnAsnValArgGlnLeuGlnArgIleThrGlyAla 739
QY 1564 GAATCATCGTCCTCTCGACCAACGCCAGAT----- 1596
Db 740 ValValLysIleProGlnGlnGlnLysArgAsnGlyGlyValAlaTyrArgHisAspAspGly 759
QY 1597 ---GAAATGAGAGATGATCGTCAGAAATTTATCGGCACTTCTTGTACGACAGTGA 1653
Db 760 LeuGlnGlnAspMetThrMetIleArgThrIleGlyAsnMetLysSerThrHisAsnVal 779
QY 1654 CAGCGCAAGATCAGGAAATTTGTACACAGTG-----AAGCAGCAG 1695
Db 780 GlnPheArgLeuAlaHisLeuValAsnGlnLysTyrTyrArgSerGlyAspHisArgAsnLys 799
QY 1696 GAGCAGAAATTAACCTCAGGAGATCGCCTGACAGCGGACGAGAT 1738
Db 800 SerSerAspTyrLysGlyGly-ArgProHisSerAlaProSer 813

RESULT 2
T19216
Hypothetical protein C12D8.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19216
R:McMurray, A.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19092
A:Accession: T19216
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-589 <MW>
A:Cross-references: EMBL:Z73969; PIDN:CAA98232.1; GSPDB:GN00023; CESP:C12D8.1a
A:Experimental source: clone C12D8
C:Genetics:
A:Gene: CESP:C12D8.1a
A:Map position: 5
A:Introns: 7/3; 25/1; 159/3; 318/3; 513/2

Alignment Scores:
Pred. No.: 8 498-11 Length: 589
Score: 274.00 Matches: 118
Percent Similarity: 41.13% Conservative: 79
Best Local Similarity: 24.63% Mismatches: 183
Query Match: 4.66% Indels: 99
DB: 2 Gaps: 18

US-09-270-437D-8 (1-3283) x T19216 (1-589)
QY 667 GTCCCAACCAAGTTGTGTGTCATCATCGAAAGAGGCGCTTGAACATTAAGAAATC 726
Db 54 IleProGlnSerAlaValGlyIleValIleGlyArgGlyGlySerGlnIleGlnGlyIle 73

```

```

QY 727 ACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAAGAACTGTGAGCTGCAGAG 786
Db 74 GlnAlaLysAlaGlyCysArgValGlnMetSerProAspAlaAspProSerSerGlyVal 93
QY 787 AAGCTGTACCATTCATCCACCCCAAGAGGACTTCTGAAGCATGCCGATGATCTT 846
Db 94 ArgMetValThrLeuGlnGlySerArgSerAsnValGlnThrAlaLysHisLeuIleAsn 113
QY 847 GAAATCATGCAGAAAAGAGCAGAT-----GAGACC 876
Db 114 GlnValValAlaArgSerGlnAsnProArgProGlnTyrGlyPheProArgAlaGlnThr 133
QY 877 AAACGACCGAAGAGATTCCTGCAAAATCTTGACACAAATGGCTTGTTGAAAGACTG 936
Db 134 ThrIleAspIleAlaIlePro-----ProAsnArgCysGlyLeuIle 147
QY 937 ATTGGAAGAAAGAGCAGAAATTGGAAGAAATTGAACATGAACAGGACCAAGATTAACA 996
Db 148 IleGlyLysSerGlyAspThrIleArgGlnLeuGlnGlyLysSerGlyCysLeuMetIle 167
QY 997 ATCTCATCTTGCAGAT---TTGACATATACAAACCCGAAAGAACATCATCATGTGAG 1053
Db 168 Leu-----ValGlnAspAsnGlnSerValSerAspGlnSerLysProLeuArgIleThr 185
QY 1054 GGCACAGTTGAGGCTGTGCCAGTGTGAGATAGAGATTATGAAGAAAGCTGCGTAGAGCC 1113
Db 186 GlyAspProGln-----LysIleGlnLeuAlaLysGlnLeuValAlaGln 200
QY 1114 TTGCAAAATGATATGCTGGCTGTTAACAACCCATCTCCGATATCTTCCAGCTGTACC 1173
Db 201 IleLeuAsnSerGlyGlyAspGlyAsnGlyGlySerGly-----LeuGlnMet 216
QY 1174 CATACACAGTTTGGCCCGTCCGCATCATCATCTTATCCAGACAGAGATGTGAT 1233
Db 217 HisHisAlaGlyGly-----GlyGlyGlyAlaSerAlaArgGlyGlnValVal--- 232
QY 1234 CTCTTCATCCCAACCCAGCTGTGCGGCCCATCATCCGGAAGAAAGAGGACACATCAAA 1293
Db 233 -----ValProArgSerSerValGlyIleIleIleGlyLysGlnGlyAspThrIleLys 250
QY 1294 CAGCTGGCAGATTCGCCGAGACCTCTATCAAGATTCCTCGGAGAGGCCAGACGTC 1353
Db 251 ArgLeuAlaMetGlnThrGlyThrLysIleGlnPheLysProAspAspProSerThr 270
QY 1354 AGCGAAGAGATGTCATCATCAACCGGAGACCCGGAACCCGATTCAGAGCCCAAGGACG 1413
Db 271 ProGlnArgCysAlaValIleMetGlyThrArgAspGlnIleTyrArgAlaThrGlnArg 290
QY 1414 ATCTTTGGAAACTGAAGAAGAAACTTCTTTAAC-----CCCAA 1455
Db 291 IleThrGlnLeuValLysLysSerThrMetGlnGlnGlyGlyGlyAsnValAlaGly 310
QY 1456 GAAAGAGTGAAGCTGCAGAGCG-----CATATCAAGATGCCCTTCCACAGCTGCG 1506
Db 311 AlMetValSerAsnGlnAlaSerThrPheTyrMetSerValProAlaAlaLysCysGly 330
QY 1507 CGGGTGTGTCGAAAGGTGGCAAGACCGTGAACGAATGCAAGAACTTGACAGTGCAGAA 1566
Db 331 LeuValIleGlyLysGlyGlnThrIleLysGlnIleAsnSerGlnSerGlyAlaHis 350
QY 1567 GTCATGCTGCTGTCGACCAAGCCGATGAAATGAGAAATGAGATGATCGTC----- 1617
Db 351 CysGlnLeuSerArgSerProThrGlyAsnAlaAspGlnLysValPheValIleLysGly 370
QY 1618 -----AGAAATTATCGGGCATCTTCTTGACCACTGCACAGCGCAAGATCAAGGAA 1671
Db 371 GlyLysArgAlaIleGlnHis-----AlaLysHisLeuIleArgIleLysValGlyAsp 388
QY 1672 ATTGTCAACAGGTGAAGCAGCAGACGACAAATACCTCAGGAGTGCCTCAACGCC 1731
Db 389 IleAlaProAsnThr----- 393
QY 1732 AGCAAGTGAAGCTCCACAGACGACAGCAAAACAGATGAATGTAGCCTTCAACAG 1791

```

```

Db 394 -----CTGACGAGATGAGACCAACGAG-----ProPheArgAsp 397
QY 1792 -----CTGACGAGATGAGACCAACGAG-----CGAGCCAGATCG----- 1827
Db 398 AsperAlaMetThrMet-----GlnThrGlnPheSerAlaProAlaGlnAsnAsn 415
QY 1828 GGACGAAACCAAAACCCATCTGAGAGATGAGAAGTCTCCGAGAGCGCGCCAGGACTTCGC 1887
Db 416 GlyGlyGlnGlnGlnTyrAsnProValAlaGlnIleProAlaAlaAlaGln----- 432
QY 1888 CGAGGCGCTGAGAAACCCGAGAGGAGGCGGAGAGGAGTTCAGCGATTTGCCAG 1947
Db 433 -----AspProTyrGlnValGlyGlyTyrGlnGlnAsnSerValTyrAlaGln 448
QY 1948 AACCAAGAGCCCGCGCTCCGCGCCAGGCGCTTTCAGAGCATTCGA 2002
Db 449 GlnThrAlaAlaProAlaAlaAlaProTyr-AlaAlaAlaGlyIleValGlnPro 466

RESULT 3
T19217
hypothetical protein C12D8.1b - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T19217
R;McMurray, A.
submitted to the EMBL Data Library, June 1996
A/Reference number: Z19092
A/Accession: T19217
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-611 <N1>
A/Cross-references: EMBL:Z73969; PTDN:CAA98233.1; GSPDB:GN00023; CESP: C12D8.1b
A/Experimental source: clone C12D8
C/Genetics:
A/Gene: CESP: C12D8.1b
A/Map position: 5
A/Intons: 181/3; 340/3; 535/2

Alignment Scores:
Pred. No.: 611
Score: 274.00 Matches: 118
Percent Similarity: 41.13% Conservative: 79
Best Local Similarity: 24.63% Mismatches: 183
Query Match: 4.66% Indels: 99
DB: 2 Gaps: 18

US-09-270-437D-8 (1-3283) x T19217 (1-611)
QY 667 GTCCCAACCCAGTGTGTCGCATCATTCGAAAGAGGCGTTGACCATPAAGACATC 726
Db 76 IleProGlnSerAlaValGlyIleValIleGlyArgGlyGlySerGlnIleGlnGlyIle 95
QY 727 ACTAAGACCCAGTCCCGGTAGATATCCATAGAAAAAGAACTGTGAGCTGCAGAG 786
Db 96 GlnAlaLysAlaGlyCysArgValGlnMetSerProAspAlaAspProSerSerGlyVal 115
QY 787 AAGCTGTACCATTCATCCACCCCAAGAGGACTTCTGAAGCATGCCGATGATCTT 846
Db 116 ArgMetValThrLeuGlnGlySerArgSerAsnValGlnThrAlaLysHisLeuIleAsn 135
QY 847 GAAATCATGCAGAAAAGAGCAGAT-----GAGACC 876
Db 136 GlnValValAlaArgSerGlnAsnProArgProGlnTyrGlyPheProArgAlaGlnThr 155
QY 877 AAACGACCGAAGAGATTCCTGCAAAATCTTGACACAAATGGCTTGTTGAAAGACTG 936
Db 156 ThrIleAspIleAlaIlePro-----ProAsnArgCysGlyLeuIle 169
QY 937 ATTGGAAGAAAGAGCAGAAATTGGAAGAAATTGAACATGAACAGGACCAAGATTAACA 996
Db 170 IleGlyLysSerGlyAspThrIleArgGlnLeuGlnGlyLysSerGlyCysLeuMetIle 189

```

```

QY 997 ATCTCATCTTTCAGAT---TTGAGCATATACACCCGAAAGACCATCTGTGAG 1053
Db 190 Leu-----ValGlnAspAsnGlnSerValSerAspGlnSerLysProLeuArgLleThr 207
QY 1054 GGCAAGTGTAGGGCTGTGCCAGTGTGAGATGAGATTTTGAAGAAAGCTGTGAGGCC 1113
Db 208 GlnAspProGln-----LysLleGlnLeuAlaLysGlnLeuValAlaGln 222
QY 1114 TTTGAAAATGATATGCTGCTGTATACACCATCCGATACTTCTCCAGCTGTACCC 1173
Db 223 IleLeuAsnSerGlyGlyAspGlyAsnGlyGlySerGly-----LeuGlnMet 238
QY 1174 CATGACCAAGTTTGCCGCTTCCCGATCATCACTTTTCCAGACAGAGATGTGAT 1233
Db 239 HsHsAlaGlyGly-----GlyGlyGlyAlaSerAlaArgGlyGlyValVal--- 254
QY 1234 CTCTTCATCCCAACCCAGCTGTGGCGCCCATCATCGGAGAAAGAGGAGGACATCAA 1293
Db 255 -----ValProArgSerSerValGlyLleLleLleGlyGlyGlnGlyAspThrLleLys 272
QY 1294 CAGTGGCGAGATTCGCGGAGCTCTTATCAAGATTGCCCTTCGCGAAGGCCACAGCTC 1353
Db 273 ArgLeuAlaMetGlnThrGlyThrLysLleGlnPheLysProAspAspProSerThr 292
QY 1354 AGCGAAAGATGTGATCATCAACCGGCGACCGGAGGCCAGTTCAAGGCCAGGAGGACG 1413
Db 293 ProGlnArgCysAlaValLleMetGlyThrArgAspGlnLleArgAlaThrGlnArg 312
QY 1414 ATCTTGGGAAACTGAAAGAGGAAACTTCTTTAAC-----CCCAAA 1455
Db 313 IleThrGlnLeuValLysLysSerThrMetGlnGlnGlyGlyGlyAsnValAlaGly 332
QY 1456 GAAGAAGTAGAGCTGGAACG-----CATATCAGAGTGCCTCTTCCACAGCTGCG 1506
Db 333 AlaMetValSerAsnGlnLysSerThrPheTyrMetSerValProAlaAlaLysCysGly 352
QY 1507 CGGGTGAATGGGAAGGTGGCAAGCCGTAAGCGAATCAGACAGCTTACACGAGTGAGAA 1566
Db 353 LeuValLleGlyLysGlyGlyGlnThrLleLysGlnLysAsnSerGlnSerGlyAlaHis 372
QY 1567 GTCATCGTCTCGTACCAACCGCCAGATGAATAATGAGAAATGATGCTC----- 1617
Db 373 CysGlnLeuSerArgAspProThrGlyAsnAlaAspGlnLysValPheValLleLysGly 392
QY 1618 -----AGAATTATCGGGCACTTCTTGTAGCCACATGTCACAGCCGAAAGTACAGGAA 1671
Db 393 GlnLysArgAlaLleGlnHis-----AlaLysHisLleuLleArgLleLysValGlyAsp 410
QY 1672 ATTGTACACAGGTGAAGCAGCAGAGCAAGAAATACCTCAGGAGTGCCTCAGAGCC 1731
Db 411 IleAlaProAsnThr----- 415
QY 1732 AGCAAGTGAAGCTCCACAGGCACACAGCAAAACAGATGAATGAGCCCTTCCACAC 1791
Db 416 -----PropheArgAsp 419
QY 1792 -----CTGACAGATGAGCAACCAAGCCAG-----CCAGCCAGATG----- 1827
Db 420 AspSerAlaMetThrMet-----GlnThrGlnPheSerAlaProAlaGlnAsnAsnPhe 437
QY 1828 GAGGAAACCAACCAACATCTGAGGATGAGATGAGAGTCTGCGAGCGCGGACGACTCTGC 1887
Db 438 GlnGlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 454
QY 1888 CGAGGCTGTGAGAACCCCAAGGAGCGGAGGAGGAGGAGGAGTCAAGCTTTGCGAG 1947
Db 455 -----AsnProTyrGlnValGlyGlyTyrGlnGlnAsnSerValTyrAlaGln 470
QY 1948 AACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2002
Db 471 GlnThrAlaAlaProAlaAlaAlaAlaProTyr-AlaAlaAlaGlyLleValGlnPro 488

```

RESULT 4

```

T49962
hypothetical protein F8M21.160 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49962
C:By: Murthy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Ler
R:By: Murthy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Ler
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24493
A:Accession: T49962
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-568 <BEV>
A:Cross-references: EMBL:AL353993; GSPDB:GN00063; ATSP:F8M21.160
C:Genetics:
A:Experimental source: cultivar Columbia; BAC clone F8M21
A:Gene: ATSP:F8M21.160
A:Map position: 5
A:Introns: 200/3; 337/3; 544/3

Alignment Scores:
Pred. No.: 1,15e-10 Length: 568
Score: 272.00 Matches: 129
Percent Similarity: 39.19% Conservative: 94
Best Local Similarity: 22.67% Mismatches: 235
Query Match: 4.62% Indels: 112
DB: 2 Gaps: 20

US-09-270-437D-8 (1-3283) x T49962 (1-568)
QY 535 CCGATGTAAGAGGTGATAGCTCCCTTCGCGCCCTCAGCGAGCCGACGATGGGAGCACTCT 594
Db 3 ProAspHisArgMet-----SerProAspHisArgAspSerHisArgLysArgSer 19
QY 595 TCCCGGAGCAAGGCGACGCGCCCTGGGGCACTTCTCAGGCGAGA----- 639
Db 20 ArgProGlnSerAspTyrAspAspAsnGlyGlySerLysArgArgTyrArgGlyAspAsp 39
QY 640 -----CAGATTGATTTCCGCTGCGGATCTGCTCCACCGAG 678
Db 40 ArgAspSerLeuValLleAspArgAspArgThrValPheArgTyrLeuLysProValLys 59
QY 679 TTTGTTGGTCCATCATCTCGGAAGAGAGGCTTGACCATTAAGAAATCACTAAGCAGAGCC 738
Db 60 LysLleGlySerValLleGlyArgGlyGlyAspLleValLysGlnLeuArgAsnAspThr 79
QY 739 CAGTCCCGGGTATATCATAGAAAGAAAGAACTGTGAGCTGACAGAGAGCTGTACCC 798
Db 80 ArgSerLysLleArgLle-----GlyGlnAlaLleProGlyCysAspGlnArgValLleThr 98
QY 799 ATCCATGCAACCCCAAGAGGAGACT-----TTGAA 828
Db 99 IleTyrSerProSerAspGlnThrAsnAlaPheGlyAspGlyGlyLysValLeuSerPro 118
QY 829 GCATGCCGATGATTTTGAATATCATGCAAGAAAGAGCAGATGAGACCAAACTAAGCCGA 888
Db 119 AlaGlnAspAlaLeuPheArgLleHisAspArgValValAlaAspAspAlaArgSerGln 138
QY 889 GAGATTGCT-----CTGAATCTTGGCAACATAGCTGTGTT 927
Db 139 AspSerProGlnGlyGlyLysGlnValThrAlaLysLeuLeuValProSerAspGlnLle 158
QY 928 GGAAGATGATTTGGAAGAAAGAGCAGAAATTTGAAGAAATTTGAACATGAAACAGGGAGC 987
Db 159 GlnCysLleLeuLysArgGlyGlyGlnLleValGlnAsnLleArgSerGlnThrGlyAla 178
QY 988 AAGATTAACAATCTCATCTTTCAGAGATTTGAGATA-----TACAACCGGAAAGAAC 1041
Db 179 GlnLleArgLleValLysAspArgAsnMetProLeuCysAlaLeuLysAsnSerAspGlnLeu 198
QY 1042 ATCACTGTGAAGGACAGTTGAGGCTGTGCAAGTGTGAGATGAGATTAATGAGAGAG 1101
Db 199 IleGlnLleSerGlyGlnValLeuLleValLysAlaLeuLeuGlnLleAlaSerArg 218

```

```

QY 1102 CTGCGTGAAGGCTTTGAAAAATGATATGCTGCTGTTAAACCCAGCTCCGATCTTCTCC 1161
    ||| |||
Db 219 LeuHsGluAsnProSerArgSerGluAsnLeuSerSerSerGlyGlyTyrProAla 238
QY 1162 AGCGCTGACCCCATCACCAGTTGGCCCG----- 1191
Db 239 GlySerLeuMetSerHisAlaGlyGlyProArgLeuValGlyLeuAlaProLeuMetGly 258
QY 1192 -----TTCGGCGCATCATCACTCT 1209
Db 259 SerTyrGlyArgAspAlaGlyAspTyrSerArgProLeuTyrGlnProProArgAsnAsp 278
QY 1210 TATCCAGAGCAGAGG---ATTGGAATCTCTTCAATCCCAACCCAGCTGGAGCCATC 1266
    ||| |||
Db 279 ProProAlaThrGluPhePheIleArgLeuValSerProValGluAsnIleAlaSerVal 298
QY 1267 ATCGGGAAGAAGGGGCGACATCAACAGCTGGCGGAGATTGGCCGAGCTTATCAAG 1326
    ||| |||
Db 299 IleGlyTyrGlyGlyAlaLeuIleAsnGlnLeuArgGlnGlnIuThrArgAlaThrIleTys 318
QY 1327 ATT-----GCCCTCGGGAAGGCCCGAGC-----GTCAAGGAAGAAGATGCTC 1368
    ||| |||
Db 319 ValAspSerSerArgThrGluGlyAsnAspCysLeuIleThrIleSerAlaArgGluVal 338
QY 1369 ATC-----ATCACCGGGCCACCGGAAGCCGATTCAGAGCCCGAGCGAGATCTTT 1419
    ||| |||
Db 339 PheGluAspAlaTyrSerProThrIleGluAlaValMetArgLeuGlnProTyrCysSer 358
QY 1420 GGGAACTGGAAGAAGGAAACTTTTAAACCCCAAGAGAAGATGAAGCTGAAGCCCAT 1479
    ||| |||
Db 359 AspTyrValGluArgAspSerGlyLeu-----ValSerPheThrThrArg 373
QY 1480 ATCAGATGCGCCCTCTTCCACAGCTGCGGGTGATGTCGAAGAAGTGGCAACCGCTAAC 1539
    ||| |||
Db 374 LeuLeuValAlaProSerSerArgIleGlyCysIleLeuGlyTyrGlyAlaIleThr 393
QY 1540 GAACTGCAAGACTTAAACAGCTCAGAAAGTC---ATCGTCTCTGTCACCAAGCCCA 1593
    ||| |||
Db 394 GluMetArgArgMetThrLysAlaAsnIleArgIleLeuGlyTyrGluAsnLeuProLys 413
QY 1594 -----GATGAAATGAGAAAGTATGCTCAGATTAATCGGCGACTTCTT 1638
    ||| |||
Db 414 ValAlaSerAspAspGluMetValGlnValAlaAsnPheMetValLeuLeuTyrPhe 433
QY 1639 GTAGCCAGACTGACACGCGCAAGATGAGGAATTTACACAGCTGGAAGCAGCAGAG 1698
    ||| |||
Db 434 SerLeuGlnPheLeuLeuSerArgLeuArgTyrPheThrTyrSerValAsnSerAsnSer 453
QY 1699 CA-----GAAATACCTTCAGAGAGTCCCTCAACAGCAGCAAGTAGAGCTCCCA 1754
    ||| |||
Db 454 AsnPheAspValGlnIleSerGlyGlnLeuAspValAlaTyrGluAlaLeuIleGlnIle 473
QY 1755 CCGCAAAACACGAGTGAATGTA----- 1778
    ||| |||
Db 474 ThrSerArgLeuArgAlaAsnValPheAspArgGluGlyAlaValSerAlaLeuMetPro 493
QY 1779 GCCCTTCCA---ACACCTGACAGATGAGCAACCAAGCAGCCAGCTCGGAGCAAA 1835
    ||| |||
Db 494 ValLeuProTyrValPro-----ValAlaProAspAlaGlyAsp-ArgPheAspTyrAs 511
QY 1836 CCAAGACCATCTGAGGAATGAGAAAGTCTGCGAGGCGCCAGGAGCTTGCAGAGCC 1895
    ||| |||
Db 511 PserArgAspSerArgArgLeuGlu-----ArgGlyAs 522
QY 1896 TGGAAACCCCAAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1955
    ||| |||
Db 522 nProTyrProGlyGlyTyrGlySerSerGly-----ValSerAlaGluGlyTyr 538
QY 1956 AGCCCGGCTCCGCGCCCGCAGGCG 1980
    ||| |||
Db 538 rSerProTyrGlyAlaProValGly 546

```

RESULT 5

```

D96554
Hypotheical protein F19C24.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96554
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso
Chiu, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96554
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-621 <STO>
A:Cross-references: GB:AE005173; NID:G11094762; PIDN:AA629695.1; GSPDB:GN00141
A:Gene: F19C24.19
A:Map position: 1
Alignment Scores:
Pred. No.: 4.56e-10 Length: 621
Score: 263.50 Matches: 129
Percent Similarity: 38.99% Conservative: 87
Best Local Similarity: 23.29% Mismatches: 209
Query Match: 4.48% Indels: 129
DB: Gaps: 25
US-09-270-437D-8 (1-3283) x D96554 (1-621)
QY 616 CCGTGGGGCACTTCTCAGCGACAGATGATTCCCGCTGGAGATCTTGTCCTACC 675
    ||| |||
Db 9 ProAlaThrThrAlaThrAlaAlaGlu---SerValHisPheArgLeuLeuCysProAla 27
QY 676 CAGTTGTTGTCGCATCATCGGAAGAAGAGGGCTTGACCAATAAGACATCATCAACGAC 735
    ||| |||
Db 28 ThrArgThrGlyAlaIleIleGlyGlyGlyGlySerValIleArgHisLeuGlnSerVal 47
QY 736 ACCCACTCCCGGTA----- 750
    ||| |||
Db 48 ThrGlySerTyrIleArgValIleAspAspIleProValProSerGluGluArgValVal 67
QY 751 -----GATATCCATTAAGAAAAGAAC 771
    ||| |||
Db 68 LeuIleIleAlaProSerGlyLysLysLysAspGluSerAsnValCysAspSerGluAsn 87
QY 772 TCTGAGCTGACAGAAAGCTCTGACCATCATGCCACCCCAAGGGGACT----- 822
    ||| |||
Db 88 ProGlySerGluGluProLysGlnGluGlySerGlySerGlyCysAlaGlyThrSerGlyGly 107
QY 823 -----TCTGAAGCATGCCGATATTTCTGAATCATGAGAAA--- 861
    ||| |||
Db 108 AspAspGluGluAlaProSerSerAlaGlnMetAlaLeuLeuArgValPheGluArgIle 127
QY 862 -----GAGGCAGATGAGAACCAACTAGCCGAAGAGATT 894
    ||| |||
Db 128 ValPheGlyAspAspAlaAlaThrValAspGlyAspGluLeuAspLysGlyGluSerGlu 147
QY 895 CCTCTG---AAATCTTGACACAATAGCTTGTTGGTGAAGATCTGGAAGAAAGAGC 951
    ||| |||
Db 148 GlyLeuCysArgMetIleValArgGlyAsnGlnValAspTyrLeuMetSerTyrGlyGly 167
QY 952 AGAATTTGAAGAAATTTGAATGAACAGGACCAAGATTAACAATTCATCTTTGAC 1011
    ||| |||
Db 168 LysMetIleGlnIleArgGluSerSerGlyAlaIleValArgIleSerSerThrAsp 187
QY 1012 GATTG-----AGCATATCAACCCGGAAGAACATCATCTGATGAAGGCAAGTTGAG 1065
    ||| |||

```

```

Db      188 GlnIleProProCysAlaPheProGlyAspValValIleGlnMetAsnGlyLysPheSer 207
QY      1066 GCCTGTGCGCACTGCTGAGATGAGATTATGAAAGAGCTGCGTGAAGCC----- 1113
Db      208 SerValIysLysAlaLeuLeuValThrAsnCysLeuGlnGlnSerGlyAlaProPro 227
QY      1114 -----TTTGAAAAATGATATGCTGGCTGTTAAACACCCGACCTCGGATAC 1155
Db      228 ThrTrpAspGlnCysProPheProGlnProGlyIleProProGlnIleThrHisSerMetLeu 247
QY      1156 TTCTCCAGCCTGTATACCCCATCACCAGTTTGCCCGTTCGCGCAT----- 1200
Db      248 TyrHisProGlnIleProAspHisProProProAsnProMetProGluAspValGlyProPhe 267
QY      1201 CATCACTCTTATCCAGAGCAGAGATTGTG--AATCTCTTATCCCAACCCAGCTGTG 1257
Db      268 AsnAspProValValGlnGlnIleValAlaPheAspLeuLeuCysProAlaAspLysVal 287
QY      1258 GCGCGCATCATCGGGAGAGGGGCGCACATCAACAACAGCTGGCGAGATTGCGCGAGCC 1317
Db      288 GlySerLeuIleGlyLysGlyValAlaValAlaGlnAlaLeuGlnAsnGlnSerGlyAla 307
QY      1318 TCTATCAAGATTGCC--CCTGCGGAAGGCCAGAGCTGACGGAAGAGATGCTATCATC 1374
Db      308 SerIleLysValSerAspProThrHis-----AspSerGlnGlnIleValIle 325
QY      1375 ACCGGGCGCACCGGAAGCCAGTTCCAG-----GCCAGAGGA-----CGGATC 1416
Db      326 SerAlaAspGlnAsnLeuGlnIleArgThrHisSerLeuLeuAlaGlnAspGlyValMetArgVal 345
QY      1417 TTGGGAAATGAAAGAGAAACTTCTTTAACCCCAAGAAAGAGTGAAGCTGGAAGCG 1476
Db      346 HisAsnArgIleValGlnIleGlyPhe-----GluProSerAlaAlaValAla 362
QY      1477 CATATCAAGTGCCTCTTCCACAGCTGCGCGGTGATTTGGAAAGGTGGCAAGCCGTG 1536
Db      363 ArgLeuLeuValHisSerProTyrIleGlyArgLeuLeuGlyLysGlyGlyHisSerLeu 382
QY      1533 AACGAACTGCAGAACTTAACTGAGTGCAGAGATC--ATCGTGCTGCTGTAACCAAGCCCA 1593
Db      383 SerGlnMetArgArgAlaThrGlyAlaSerIleAspValPheAlaAspSerGlnAlaThr 402
QY      1594 -----GATGAAATGAGAGAGTATCTGCAAGATTAATGAGGCACTTCTTGTGACCCAG 1647
Db      403 LysIleTyrGlnSerGlnHisAspGlnIleValGlnValIleGlyAsnLeu-----Lys 419
QY      1648 ACTGCACACCGCAGAGATCGAGGAATTGTACACACAGGTGAGCAGCAGAGAAATAC 1707
Db      420 ThrValGlnAspAlaLeuPheGlnIleLeuCysArgLeuArg-----GlnAlaMetPhe 437
QY      1708 CCTCAGGAGTGGCTTCACAGCGCACAGAGTGGCTCCACAGGACAGCAAGCAAAACAAC 1767
Db      438 Pro-Gly--ArgLeuProPheGlnGlyMetGlyGlyProProProPheMetGly- 455
QY      1768 GGATGATGATAGCCCTTCCAAACACTGACAGAGATGAGCCAAACGACGACGACAGATCG 1827
Db      456 -----ProTyrProGlnProProProProPheGlyPro----- 466
QY      1828 GGAGCAAAACCAAGCAACATCTGAGAAATGAAAGTCTGCGGAGCGCGCCAGGACTCTGC 1887
Db      467 -----ArgGlnTyrProAlaSerProA 474
QY      1888 CGAGGCGCTGAGAACCCCGAGGGCGG-----AGAGGGGGCGGGAGAGTTCAGCAGAGTT 1941
Db      474 spArgTyrHisSerProValGlyProPheHisGlnIleArgHisCysHisGly--ProGlyP 493
QY      1942 TGGCAGAAACCAAGGACCGCCGCC--TCCCGCCCCCAGGGGCTTTCGAGGCTTTCAGCCA 1998
Db      493 heAspArgProProGlyProGlyPheAspArgProPro----- 505
QY      1999 TCACATTCACCAATCCACTGGATCTCTCTGAATCTCC 2036
Db      506 -----SerProMetSerTrpThrPro 512

```

```

RESULT 6
A53184
myc far upstream element-binding protein - human
N:Alternate names: FUSE-binding protein
M:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text_change 05-Nov-1995
C:Accession: A53184
R:Duncan, R.; Bazar, L.; Michelotti, G.; Tomonaga, T.; Krutzsch, H.; Avigan, M.; Levens,
  Genes Dev. 8, 465-480, 1994
A:Title: A sequence-specific, single-strand binding protein activates the far upstream e.
A:Reference number: A53184; MUID:94170991; PMID:8125259
A:Accession: A53184
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-644 <DUN>
A:Cross-references: GB:U05040; NID:9460151; PID:AAA17976.1; PUD:9460152
C:Keywords: DNA binding

Alignment Scores:
Pred. No.: 5.4e-10 Length: 644
Score: 262.50 Matches: 140
Percent Similarity: 37.73% Conservative: 86
Best Local Similarity: 23.37% Mismatches: 237
Query Match: 4.468 Indels: 136
DB: Gaps: 25

US-09-270-437D-8 (1-3283) x A53184 (1-644)
QY      336 CCCTCTCACTCGCAGCTGGAGGAGTGTGGATGAGACTTTTGCTCAATATGGAGACGTGA 395
Db      10 ProSerSerGlySerLeuGlyGlyGlyGly-----GlyGlyGly 22
QY      396 GAATGTGAACAAGTCAACACACAGACAGAAACCCCGTTGTCAAGTCAATATGCAAC 455
Db      23 GlyGlyGly-GlyValAsnAspAlaPheLysAspAlaLeuGlnArg-----AlaAsp 39
QY      456 AAGAGAAAGCAAAATATGCCATGAGAGAGCTTAACCGGGCATCACTTGAAGACATATC 515
Db      39 GlnIleAlaAlaLysIleGlyLysAspAlaGlyThrSerLeuAsnSerAsnAspLysGln 59
QY      516 CTTCAAGATTCTCTCATCTCCGAGATGAGAGTGAAGTCCCTGACCCCTCAGCGAGC 575
Db      59 YTrGlyGlyGlnLysArgProLeuGlnAspLysAspGlnProAspAlaLysValAla 79
QY      576 CCAGCGTGGGACCACTCTTCCCGGAGCAAGCCACGCCCTGGGGGCACTTTCAGGC 635
Db      79 ArgGlnAsnAspSerPheGlyThrGlnLeu-----ProProMetHisGlnGlnIle 97
QY      636 CAGACAGATTGATTTCCGCGTGGATCTGTGCTCCACCCAGTTTGTGGTCCATCAT 695
Db      97 rArgSerVal--MetThrGlnGlnIleLysValProAspGlyMetValGlyPheIleI 116
QY      696 CGGAAAGAGGGCTTGACCAATAAGAAATCACTAAGACAGACCCAGCTCCGGGTAGATAT 755
Db      116 eGlyArgGlyGlyGlnIleSerArgIleGlnGlnGlnSerGlyCysLysIleGlnI 136
QY      116 eGlyArgGlyGlyGlnIleSerArgIleGlnGlnGlnSerGlyCysLysIleGlnI 136
Db      756 CCATGAAAGAGAACTTGAAGCTGAGCTGCAAGAAAGCTGTACCATTCATGCCAACCCCA 815
Db      136 e--AlaProAspSerGlyLysLeuProGlnIleArgSerCysMetLeuThrGlyThrProG 155
QY      816 GGGGACTTGTGAAGCATGCCGATGATTTCTTGAATATCAGCAAA----- 861
Db      155 uSerValGlnSerAlaLysArgLeuLeuAspGlnIleValGlnLysGlyArgProAlaP 175
QY      862 -----GAGGCGATGAGACCAAACTAGCCGGAAGAGATTCCTGTGAATCTTGGC 911
Db      175 oGlyPheHisHisGlyAspGlyProGlyAsnAlaValGln-----GlnLeuMetI 192
QY      912 ACACAAATGCTGTGTGGAAGACTGATTTGAAAGAGGAGGAGAAATTTGAAGAAATTTGA 971
Db      192 eProAlaSerLysAlaGlyLeuValIleGlyLysGlyGlyGlnThrIleLysGlnLeuG 212

```

```

QY 972 ACATGAACAGGAGCAAGATACATCTTTCAGAGATTGACATATACACC 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 ngluargalaglyVallysMetValMet-----lleInAsp-----GlyPr 226
QY 1032 GGAAGAAGACC-----ATCACTGTGAAGGCGACAGTTGAGCGCTGTGCCAG 1076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 226 oglnmsnThrIglYalaAspLysProleuArglleThrIglYAspProtyrLysValIngl 246
QY 1077 TGCTGAGATAGATATTAGAAAGAGCTGTCGAGGCTTTGAAATGATATGCTGCNGT 1136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 nalalysglMetValleuGluIleuIleArgAsp----- 257
QY 1137 TAAACACCCACTCCGACTTCTTCAGGCTGTAACCCCATGACAGTTGGCCGCTGCC 1196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 -----GlnIglYlPheArgGluVal-----ArgsnIglYrGly----- 269
QY 1197 GCATCATCACTCTTATCCAGAGCAGAGATTTGATCTTTCATCCCAACCCAGCTGT 1256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 -----SerArglleIglYlPheArgsnIglYlLeaspValProIleProArgPheAlaVa 287
QY 1257 GAGGCGCATCATCGGGAAGAAGGGGGGCAACATCAACAGCTGGCGAGATTGCCGAGC 1316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 287 lglYlleVallleIglYlPheArgsnIglYlMetlleLysLyslleGlnsnAspAlaGlyVa 307
QY 1317 CTCTATCAAGATTGCCCTGCGAAGGCGCCAGACGTACGCAAGAGATGTCATCATCAC 1376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 lArglleGlnPheLysProAspArgGly---ThrThrProIluArglleAlaGlnlleTh 326
QY 1377 CGGGCCACCGGAA---GCCCACTTCAGAGCCGAC----- 1407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 rGlyProProAspArgCysGlnHisAlaAlaGlnlleIleThrAspLeuArgSerVa 346
QY 1408 -----GAGCGATCTTGGGAA----- 1425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 lGlnAlaGlyAsnProIglYlProIglYlArgGlyArgGlyArgGlyIngl 366
QY 1426 -----CTGAAGAAGAAACTTCTTAAACCCCAAGA 1457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 yAsnTrpAsnMetGlyProProIglYlLeuGlnIglPheAsnIle----- 382
QY 1458 AGAAGTGAAGCTGGAAGCCATATCAGATGCCCTCTTCAACAGCTGGCGGATGTGG 1517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 -----ValProThrGlyLysThrIglYleuIlelleIgl 393
QY 1518 CAAGGTGCAAGACCGTGAACGAATGCAAGACTTACCAAGTGCACAAATCATCGGCC 1577
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 393 yLysGlyGlyGlnThrIleLysSerlleSerGlnInserGlyAlaArglleGluIngl 413
QY 1578 TCGTGAACAAACCCGATGAAATGAAGAAGT---ATCGTCAATATTCGGGAC-- 1632
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 413 nArgAsnProProProAsnAlaAspProAsnMetLysleuPheThrIleArgGlyThrPr 433
QY 1633 -----TTCTTGCTAGCCAGACTGCACAGCCGACAGATCAAGGAAATTTGACACA 1682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 433 ogInGlnIleAspIlyAlaArgIleuIleGlnGluLyslleGlyProValAsnPr 453
QY 1683 GGTGAAGACAGCAGAGACAATACCTCGAGAGTGCCTCACAGGCGACGAAGGAG 1742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 453 oleuGlyProProValProHisGlyProHisGlyVal-ProGlyProHisGlyProProG 473
QY 1743 CTCGCAAGAGCAGCAAAACAAACGATGATGTAGCCCTTCAACACCTGCAGAAAT 1802
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 lYProProGly-ProGly-----ThrPro-----Met 481
QY 1803 AGACCAAGCAGCAGCCAGCATCGGAGCAACAAAGCAATCTGAGGAATGAGAAGT 1862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 GlyProIlyAsnProAlaProIlyAsnProGlyPro----- 493
QY 1863 CTGCGAGGCGGCGAGGACTGCGGAGG---CCCTGAGAAACCCAGGGCGGAGAGG 1919
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 -----ProGlyProAlaProHisGlyProProAlaProIlyAlaProGlnGly 509
QY 1920 GGGGGGAAAGTCAAGCCAGGTTTGCAGAA-----CAACGAGCCCCGCC 1964

```

```

Db 510 TrpGlyAsnAlaIlyrProHisTrpGlnGlnAlaProProAspProAla 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 7
T27855
Hypotheical protein ZK418.9 - Caenorhabditis elegans
CSpecies: Caenorhabditis elegans
CDate: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
CAccession: T27855
R,Fulton, U.
Submitted to the EMBL Data Library, April 1994
A:Description: The sequence of C. elegans cosmid ZK418.
A:Reference number: Z20430
A:Accession: T27855
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-768 <FNU>
A:Cross-references: EMBL:U00047; PIDN:AAA50693.1; CESP:ZK418.9
A:Experimental source: strain Bristol N2
C:GeneticS:
A:Gene: CESP:ZK418.9
A:Introns: 166/1; 221/3; 243/1; 279/1; 387/3; 490/1; 567/3; 727/2

Alignment Scores:
Pred. No.: 6,67e-10 Length: 768
Score: 261.50 Matches: 129
Percent Similarity: 39.92% Conservative: 83
Best Local Similarity: 24.29% Mismatches: 226
Query Match: 4.44% Indels: 93
DB: 2 Gaps: 20

US-09-270-437d-8 (1-3283) x T27855 (1-768)
QY 560 CGCCCCCTC---AGCGAGCCGAGGTGGGAGCACTTCGCCGAGCAAGCCACGCC 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 ArgProleuAspSerGlnIleLeuAspGlyAspLeuIleProThrIlyLysSerSerGlu 253
QY 617 CTGGGGGCACTT---CTGAGCGCAACAGATTGATTTCCGCTGGGATCTGTCCCA 673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 254 ValGlyAspLeuAsnMet-GlyAspSerAspLyslleThrAspIleTyPro-ValProG 273
QY 674 CCAAGTTGTGTGGTCATATCGCAAGAGAGAGGCTTGACCATTAAGAAATCACTAAG 733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 luisValValGlyLeuVallleGlyLysGlyLysGlnIleArgllelleGlnIntr 293
QY 734 AGACCAAGTCCGAGTATATCATGAAGAAAGAACTCTGAGAGTGCAGAGAGAGCTG 793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 hsrGlyCysArgValGlnMetAspProAspHisGlnSerValAsnGlyPheArgAsn 313
QY 794 TCACCATTCATGCCACCCGAGAGGAGCTTTCAGAGCATGCCGATGATTTTGAATCA 853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 ysrThrIleGlnGlyProProAspGlnValAlaValAlaArgGlnMetlleThrGlnValI 333
QY 854 TG-----CAGAAAGAGCAGATGAGCAACAACTAGCCGAGAGATTCCTGAAA 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 333 leAsnArgsnGlnThrIglYAlaGlnProGlyAlaAlaProGlyGluValThrGluInl 353
QY 905 TCTTGCACACATGCTGTTGGAAGACTGATTTGAAAGAAAGCAAGCAAAATTTGAGA 964
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 etleuIleProAlaAspLyslleGlyLeuVallleGlyLysGlyGlnThrIleArgI 373
QY 965 AAATTGAACATGAACAGGAGCCAGATTAACAATTCATCTTTCAGAGATTGAGCATAT 1024
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 373 leValGlnGlnInserGly-----leuArgAsnCysAsnValVal 386
QY 1025 ACAACCCGGAAGAACATCACTGTGAAGGCAACATTTGAGGCTGT-----G 1072
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 al-----GlnGluThrThrThrAlaThrGlyGlnProLysProleuArgMetlleGlyS 404
QY 1073 CCAAGTGAAGATAGATTATGAAGACCTGCGTGAAGGCTTTGAAATGATATCTGG 1132
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 404 exProAlaAlaIleGlnThrAlaLysAlaLeuValHisAsnIleMetAsn----- 420

```



```

Db      69  LeuArgMetLeuCysLeuValIleuHisAlaSerLeuIleValIleGlyIleGlyAlaThr 88
QY      715  ATAAAGAAATCATAGTACAGACCCAGTCCGGGAGATTCATCCATAGAAAAGAAACCT 774
Db      89  IleSerArgIleIleYserSerGluThrSerAlaArgIleAsnIleSerAsnIleArg--- 107
QY      775  GGAGGTGCAGAGAGACCTGTCACCATCCATGACCCAGAGGGGAGCTTGTGAAGCATGC 834
Db      108  GluValProGluArgIleValIleValArgGlyThrCysAspAspValAlaIleAlaIle 127
QY      835  CGCATGATTTCTTGAATATCATGACAGAAAGAGCGAGATGAGACCAACTAGCCGAA----- 888
Db      128  GlyMetIleValArgAlaIleuGluIleuGlnIleGlyAsnGluAspAsnGlyGluAspIle 147
QY      889  GAGATTTCTTGAATATCTTGGCACAACATGCTGTTGTGAAGCATGATGGAAGAAAGA 948
Db      148  GluIleSerIleAsnLeuIleuIleProHisIleuMetGlyCysIleIleIleIleYsArg 167
QY      949  GGCAGAAATTTGAGAAATTTGAAATGAAATGAAACAGGAGACCAAGATTAACAATCTTGTG 1008
Db      168  GlySerArgIleuArgGluIleGluAspIleuSerAlaIleAlaIleuPheAlaSerProAsn 187
QY      1009  CAGGATTTGAGCAATATACACCCGAAAGAACATCATCTGTGAAGGCGACAGTGAAGGCC 1068
Db      188  Gln-----LeuLeuIleuSerAsnAspArgIleLeuThrIleAsnGlyValProAspAla 205
QY      1069  TGTGCACATGCTGAGATGAGATGATTAAGAAAGCTGCGTGAAGCCCTTGAATATGATATG 1128
Db      206  IleHisIleAlaThrPheThrIleSerGlnThrIleuAsn---PheGlnMetGluSer 224
QY      1129  CTGGCTGTAAACACC---CACTCCGAGATCTTC-----TCCAGCCTG 1167
Db      225  ProGlnIleAsnValIleYsArgSerIleThrGlnProThrGlnPheAsnSerValIleu 244
QY      1168  TAACCCCATCACCAAGTTGGCCCGTCCCGCAT-----CATCACTCTTATCCAGAG 1218
Db      245  IleAspHisSerGlnProAsnThrIlePheHisGlnArgAsnHisGlnIleThrIleProSer 264
QY      1219  CAGGAGATTTGTG-----AATCTTCATCCCAACCAAGCTGTGGCGGCATC 1266
Db      265  AspIleYsLeuIleuSerThrIleYsProAsnIleAsnLeuProIleSerThr----- 281
QY      1267  ATCGGAGAGAGAGGGCGACACATCAAAACAGCTGGCG-----AGATTCGCGGAGCTCTT 1320
Db      282  -----LeuLeuSerMetAlaThrProGlnIleThrIleAlaSer 294
QY      1321  ATCAAGATTTGCCCTTCGGAAGGCCCAAGAC---GTACAGGAAAGATGTCATCATCACCC 1377
Db      295  ValAlaAsnAlaThrIlePheGlnProAsnPheValIleProAsnValIleThrValIleuAsp 314
QY      1378  GGCGCCACCGGAGAGCCGATTCAGGCCAGGATCTTGGGAAACTGAAAGAGAGAA 1437
Db      315  GlyProValIleIleSerProAlaProGlyAsnHisIleuLeuMetAsnIleValGlnGlnIle 334
QY      1438  AACTTCTTTAAACCCCAAGAAAGAAAGTGAAGCTGAAAGCGCATATCAAGAGTCCCTTCC 1497
Db      335  IlePheIleAspIleIleYs----- 340
QY      1498  ACAAGTGGCGGGGTGATTTGGCAAAAGGTGGCAAGACCGTGAACCAATGCAAGACTTAACC 1557
Db      341  PheValGlyAsnValIleGlyIleYsAspGlyIleHisIleAsnSerValIleYsGluSerThr 360
QY      1558  AGTGCAGAAAGTACGTGCTGCTGACCAAAAGCCAGATGAA---AATGAGGAAGTGATC 1614
Db      361  GlyCysSerIleIleIleIle-----GlnAspProValGluGlySerSerGluAspArg 377
QY      1615  GTCAAGATTAATCGGCACTTTCTTGTAGCCAGACTGCA-----CAGCGCAG 1662
Db      378  LeuThrIleArgGlyThrPheMetAlaSerGlnAlaIleIleMetLeuIleSerAsnIleYs 397
QY      1663  ATCAAGGAAATTTGTACAAACAGGTGAAGCAGCAGAGCAAGAAATACCTT 1710
Db      398  Ile---GluIleAspArgSerAsnAlaGluArgIleYsArgIleSerPro 412

```

RESULT 13

S58529

Alpha-complex protein 1 - human

NAlternate names: nucleic acid-binding protein; protein PCBP-1

CSpecies: Homo sapiens (man)

CDate: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 01-Dec-2000

CAccession: S58529; S58523; S5678; S43489; S41378; S42472

R.Kiledjian, M.; Wang, X.; Liebhafner, S.A.

EMBO J. 14, 4357-4364, 1995

A>Title: Identification of two KH domain proteins in the alpha-globin mRNA stability com

A.Reference number: S58523; MUID:96016208; PMID:7556077

A.Accession: S58529

A>Status: preliminary

A.Molecule type: protein

A.Residues: 1-356 <K1P>

A.Accession: S58523

A>Status: preliminary

A.Molecule type: protein

A.Residues: 125-139;251-265;315-323 <K12>

R.Lieffers, H.; Deigaard, K.; Celis, J.E.

Eur. J. Biochem. 230, 447-453, 1995

A>Title: Characterisation of two major cellular poly(UC)-binding human proteins, each co

A.Reference number: S5678; MUID:95331278; PMID:7607214

A.Accession: S5678

A>Status: nucleic acid sequence not shown

A.Molecule type: mRNA

A.Residues: 1-204,'V',206-356 <LEF>

A.Cross-references: EMBL:X78137; NID:G460770; PIDN:CAA55016.1; PID:G460771

A.Experimental source: ANA cells (transformed human amnion cells)

A.Note: Submitted to the EMBL Data Library, March 1994

R.Aasheln, H.C.; Loukianova, T.; Deggerdal, A.; Smeland, E.B.

Nucleic Acids Res. 22, 959-964, 1994

A>Title: Tissue specific expression and cDNA structure of a human transcript encoding a

A.Reference number: S43489; MUID:94203810; PMID:9152927

A.Accession: S43489

A>Status: translation not shown

A.Molecule type: mRNA

A.Residues: 1-298,'H' <AAS>

A.Cross-references: EMBL:Z29505; NID:G444020; PIDN:CAA82631.1; PID:G444021

C.Genetics:

A.Gene: GDB:HNRPX

A.Cross-references: GDB:344947

C.Keywords: RNA binding

Alignment Scores:

Pred. No.:	1.73e-07	Length:	356
Score:	225.00	Matches:	82
Percent Similarity:	40.40%	Conservative:	61
Best Local Similarity:	23.16%	Mismatches:	123
Query Match:	3.82%	Indels:	88
		Gaps:	11

US-09-270-437D-8 (1-3283) x S58529 (1-356)

```

QY      619  GGGGAGACTTCTCAGGACGACAGATTTCCGCTGCGATTCGTCGATCCGACCCAG 678
Db      4  GlyValThrGluSerGlyLeuSerValThrLeuThrIleArgLeuLeuMetHisGlyIle 23
QY      679  TTGTTGGTGCATCATCGGAAAGAGAGGCTTGACCATTAAGAAATCATCAAGACGACC 738
Db      24  GluValGlySerIleIleGlyIleYsIleGlyIleSerValIleYsArgIleArgGluIle 43
QY      739  CAGTCCGGGTATATTCATTAAGAAAGAGACTCTGAGCTTCAGAGAAAGCTGTACCC 798
Db      44  GlyAlaArgIleAsnIle-----SerGluGlyAsnIleYsProGluArgIleIleThr 60
QY      799  ATCAATGCAACCCCAAGAGGAGACTTCTGAAGATGCGCATGATCTTGAATATCAGCAG 858
Db      61  LeuThrGlyProThrAsnAlaIlePheIleValAlaPheAlaIleIleIleAspIleGlu 80
QY      859  AAGAGGACATGAGACCAACTA-----GCCGAGAGATTCCTCTG 900

```

```

Db      81 GluAspIleAsnSerSerMetThrAsnSerThrAlaIleSerArgProProValThrLeu 100
QY      901 AAAATCTTGACACAAATGGCTTGGTAGAGACTTGTAAGAAAGAGCGAAATTTG 960
Db      101 ArgLeuValValProAlaThrGlnCysGlySerLeuIleGlyLysGlyCysLysIle 120
QY      961 AAGAAATTTGAACATGAAACAGAGACCAAGATACAGATCTCATCTTGGCAGATTGAGC 1020
Db      121 LysGluIleArgGluSerThrGlyAlaGlnValGlnValAlaGly-----AspMetLeu 138
QY      1021 ATATCAACCCGGAAGAAAGCAATCTGCTGTAAGGCGACAGTTGAGCGCTTGCCAGTCT 1080
Db      139 ProAsnSerThrGluArgAlaIleThrIleAlaGlyValProGlnSerValThrGlnCys 158
QY      1081 GAGATGAGATT-----ATGAAGAAGCTGCGTAGGCGCTTGGAAGAAATATG 1128
Db      159 ValLysGlnIleCysLeuValMetLeuGluThrLeuSerGlnSerProGlnGlyArgVal 178
QY      1129 CTGGCTGTTAACACCCACTCCGGAATCTTCCAGCCTG----- 1167
Db      179 MetThrIleProTyrGlnProMetProAlaSerSerProValIleCysAlaGlyGln 198
QY      1168 -----TACCCCAT----- 1176
Db      199 AspArgCysSerAspAlaIleGlyTyrProHisAlaThrHisAspLeuGluGlyProPro 218
QY      1177 -----CACCAAGTTGGCCG----- 1191
Db      219 LeuAspAlaTyrSerIleGlnGlyGlnHisThrIleSerProLeuAspLeuAlaLysLeu 238
QY      1192 -----TTCCGCATCATC----- 1206
Db      239 AsnGlnValAlaArgGlnGlnSerHisSerHisMetMetHisGlyGlyThrGlyPheAla 258
QY      1207 -----TCTTATCCAG----- 1218
Db      259 GlyIleAspSerSerSerProGluValLysGlyTyrTrpAlaSerLeuAspAlaSerThr 278
QY      1219 CAGAGATTTGATCTTCTTATCCCAACCCAGGCTGCGGCGCATTCATTCAGAAAGAG 1278
Db      279 GlnThrThrHisGluLeuThrIleProAsnAsnLeuIleGlyCysIleIleGlyArgGln 298
QY      1279 GGGGACACATCAACAGCTGCGAGATTCGCGAGAGCTCTATCAAGATTGCC---CCT 1335
Db      299 GlyAlaAsnIleAsnGlnIleArgGlnMetSerGlyAlaGlnIleLysIleAlaAsnPro 318
QY      1336 GCGGAGGCGCCAGACGTCAGCGAAAGATGCTCATCATCCGCGCCACCGGAAGCCGAG 1395
Db      319 ValGluGlySer-----SerGlyArgGlnValThrIleThrGlySerAlaIleSerIle 336
QY      1396 TTCAAGGCCGAGGAGCGATTTGGGAAACTGAAAGAGAA 1437
Db      337 SerLeuAlaGlnTyrLeuIleAsnAlaArgLeuSerGln 350

```

RESULT 14

T04533

hypothetical protein F28U12.30 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 14-May-1999

C:Accession: T04533

R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; submitted to the Protein Sequence Database, February 1998

A:Reference number: 215377

A:Accession: T04533

A:Molecule type: DNA

A:Residues: 1-846 <BEV>

A:Cross-references: EMBL:AL021710

A:Experimental source: cultivar Columbia; BAC clone F28U12

C:Genetics:

A:Map position: 4

A:Insertions: 94/3; 140/3; 161/3; 191/3; 232/3; 262/1; 287/1; 293/3; 300/3; 503/3; 675/3; 7

A:Note: F28U12.30

Alignment Scores:

Pred. No.:	2,62e-07	Length:	846
Score:	224.00	Matches:	117
Percent Similarity:	33.61%	Conservative:	87
Best Local Similarity:	19.28%	Mismatches:	198
Query Match:	3.81%	Indels:	205
DB:	2	Gaps:	20

US-09-270-437D-8 (1-3283) x T04533 (1-846)

```

QY      370 CTTTGGCTCAATATGGAGAGTAGTGAAAGATGCAACA----- 417
Db      268 LeuLeuAspSerTyrGly-----HisThrIleGlyValAsnThrAlaThrPheThr 284
QY      418 -----GACACAGAAACCGCGTTGTCAACGTC 444
Db      285 ArgLysGlySerGlyMetSerSerGlyLeuIleGluThrAspSerGluValLeuAsnSer 304
QY      445 ACATATGCAACAAGAGAAAGCAAAATATGACATGAGAAAGCTAAGCGGCATCAGTTT 504
Db      305 ValTyrLeuThrMetValGluArgLys----- 313
QY      505 GAGAACTACTCTTCAAGATTCTTACATCCCGGATGAAGAGTGAGCTCCCTTGCGCC 564
Db      314 -----Lys 314
QY      565 CCTCAGGAGGCCACCGCTGGGAGCACTTTCGCGGAGCAAGCCACGCCCTGGGGGC 624
Db      315 ArgLysGlnIleGlnArgAsnAsnSerGluSerAsnArg----- 327
QY      625 ACTTCTCAGGCGACACAGATGATTCCG-----CTG 657
Db      328 ---AsnGlnLysArgGlyLeuSerHisAspLysIleAsnArgAspGluLeuValValTyr 346
QY      658 CGGATCCGTGTCGCCACCCAGTTTGTGTCATCCGGAAGAGAGGCTTGACCATTA 717
Db      347 ArgIleLeuCysProIleAspValValGlyValIleGlyLysSerGlyLysValIle 366
QY      718 AAGAACTACTTAAGCAGACCCAGTCCGGGTAGATATCCATAGAAAGAAAGACTGGA 777
Db      367 AsnAlaIleArgHisAsnThrIleValAlaLysIleLysValPhe---AspGlnLeuHisGly 385
QY      778 GCTGCAGAGAGAGCTGATCCATCATCCACCCCAAGAGGAGACTTGAAGCATGCCGC 837
Db      386 CysSerGlnArgValIleThrIleTyrCysSerValLysGluLysGlnGlu----- 403
QY      838 ATGATTCTGAAATCATCAGAAAGAGCGAGATGAGCAAACTAGCCGAAGATTCTT 897
Db      404 -----IleGlyPheThrLysSerGluAsnGluProLeuCysCysAlaGlnAspAlaLeu 421
QY      898 CTGAAATCTTGGCACCAATGCGTTGGT----- 927
Db      422 LeuLysVal-----TyrAspAlaIleValAlaSerAspGluGluAsnThrLysThr 439
QY      928 -----GGA 930
Db      440 AsnValAspArgAspAspAsnLysGluCysArgLeuLeuValProPheSerGlnSerSer 459
QY      931 AGACTGATTTGAAAGAGCGCAAGAAATTTGAAGAAATTTGAACATGAACAGGAGCAAG 990
Db      460 SerLeuIleGlyLysAlaGlyGluAsnIleLysArgIleArgArgAlaGlnAlaSer 479
QY      991 ATAACTATCTCATCTTGGCAGATTGAGC----- 1020
Db      480 ValLysValValSer---LysAspValSerAspProSerHisValCysAlaMetGluTyr 498
QY      1020 ----- 1020
Db      499 AspAsnValValIleSerGlyGluProGlnSerValLysGlnAlaLeuPheAlaVal 518
QY      1021 -----ATATAC-----AACCGGAAGAACATCATCTGTAAGGCGCAGTTGAG 1065
Db      519 SerAlaIleMetCyrLysIleAsnProArgGluAsnIleProLeuAspSerThrSerGln 538

```

```

QY 1066 GCCTGTGCGACGTGTGAGATGAGATTGAGAAAGCTGCGTGAAGCC----- 1113
Db 539 AspVal1ProAlaIleSerVal1IleVal1ProSerAspLeuSerAsnSerVal1TyrProGln 558
QY 1114 -----TTTGAAGAAAT----- 1122
Db 559 ThrGlyPheTyrSerAsnGlnAspHisIleLeuGlnGlnGlyAlaGlyValProSerTyr 578
QY 1123 ---GATATGCTGGCTGTTAAACCCACCTCCGATAC----- 1155
Db 579 PheAsnAlaLeuSerValSerAspPheGlnGlyTyrAlaGluThrAlaAlaAsnProVal 598
QY 1156 -----TTTCCAGCTGTGATCCCATCATCAGCTTTGGCCCTCCCGCATCATCAG 1206
Db 599 ProValPheAlaSerSerLeuProValThrHisGlyPheGly-----GlySer 614
QY 1207 TCTTATCCAGAGACGAGATGTGATCTTCATCCCAACCCAGCTGTGGCCCATC 1266
Db 615 SerArgSerGluGluLeuValPheLeuValLeuCyProLeuCyAsnIleMetArgVal 634
QY 1267 ATCGGAGAGAGAGGGGACACATCAACAGCTGCGAGATTGCGGAGCTCTTATCAG 1326
Db 635 IleGlyLysGlyGlySerThrIleLysArgIleArgGluAlaSerGlySerCysIleGlu 654
QY 1327 ATTGCCCCCGGAGAGGCCACACGTCAGCGGAAAGATGTCATCATCAGCGGCGACCG 1386
Db 655 ValAsnAspSerArgThrLysCysGlyAspAspGluCysValIleLeuValThrAlaThr 674
QY 1387 GAAGCCCATGTTCAAGGCCGAGGACGAGATCTT----- 1419
Db 675 GluIleLeuPheCysCysLeuSerThrProPheValPheMetGlnSerProAspMet 694
QY 1420 -----GGAAACTGAAAGAGAAACTTCTTAAACCCCAAGAGAA 1461
Db 695 LysSerMetAlaValGluAlaValLeuLeuLeuGlnGluTyrIleAsnAspGluAspAla 714
QY 1462 GTGAGAGCTGGAAGCGCATATCAGAGTCCCTCTCCACAGCTGGCGGGTGAATGGCAG 1521
Db 715 GluAsnValIleSerGlnLeuValSerLysValIleGlyCysValIleGlyLys 734
QY 1522 GGTGGCAGACCGGTGACAGACTGCAAGAACTTAACGATGTCAGAGATCATCTGCTCGT 1581
Db 735 SerGlySerValIleLeuGlnIleArgLysArgThrAsnAlaSerIleCysIleSerLys 754
QY 1582 GACCAAAACCCGACATGAAAAATGAGAGTGCATCATTAATTCGGGCACTTTTGGCT 1641
Db 755 GlyLysLysAspAsp-----LeuValGluValSerGlyGluVal----- 767
QY 1642 AGCCACACTGCAACGCGCAAGATCAGGAATGTACAAAGGTGAGAGCAGCAGGACAG 1701
Db 768 -----SerSerValArgAspAlaLeuIleGlnIleValLeuArgLeuVal 782
QY 1702 AATATCCCTCAGAGGATCGCTCAACGCGCAGCAAGTGAAGCTCCACAGGACCAACAA 1761
Db 782 ggluAspValIleuGlyAspLysAspSerValAlaIleThrArgLysProProAla----- 799
QY 1762 AACACGAGATGAATGTAGC 1780
Db 800 -ArgThrAspAsnCysSer 805

RESULT 15
A44125
high density lipoprotein-binding protein, 110K - human
C:Species: Homo sapiens (man)
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
C:Accession: A44125
R:McKnight, G.L.; Reasoner, J.; Gilbert, T.; Sundquist, K.O.; Hokland, B.; McKernan, P.A.
U. Biol. Chem. 267, 12131-12141, 1992
A>Title: Cloning and expression of a cellular high density lipoprotein-binding protein
A:Reference number: A44125; MIMD:92291094; EMD:1318310
A:Accession: A44125
A>Status: preliminary

```

```

A:Molecule type: mRNA
A:Residues: 1-1268 <MKG>
A:Cross-references: GB:M64098; GB:M83789; NID:G183891; PIDN:AAA35962.1; PID:G183892
A>Note: Sequence extracted from NCBI backbone (NCBIN:106862, NCBIPI:106863)
C:Superfamily: vlg11m

Alignment Scores:
Pred. No.: 3,46e-07 Length: 1268
Score: 223.00 Matches: 104
Percent Similarity: 38.79% Conservative: 69
Best Local Similarity: 23.32% Mismatches: 173
Query Match: 3.79% Indels: 100
DB: 2 Gaps: 18

US-09-270-437d-8 (1-3283) x A44125 (1-1268)

QY 460 GAGAGAGAAAATATGCCATGAGAGAGTGAAGCGG----- 495
Db 445 GluAspValAsnValAlaGlnGlnGlnIleGlnGlyMetValLysAspLeuIleAsnArg 434
QY 496 -----CATCAGTTTGAAGACTAC----- 513
Db 435 MetAspTyrValGluIleAsnIleAspHisLysPheHisArgHisLeuIleGlyLysSer 454
QY 514 TCCTTCAAGATTTCCTAATCCCGAGTGA--GAGGTAGCTCCCTTCGCCCTCCAG 570
Db 455 GlyAlaAsnIleAsnArgIleLysAspGlnTyrLysValSerValArgIleProProAsp 474
QY 571 CGAGCCACGCGGGGACCATCTTCCCGGAGAGCAAGCGCCCTGGGGCACTTCT 630
Db 475 SerGluLys-----SerAsnLeuIleArgIleGlyGlyAsp---ProGlnGlyValGln 491
QY 631 CAGGCCAGACAG-----ATTGATTTCCCGCTGCGG----- 660
Db 492 GlnAlaLysArgGluLeuLeuGluLeuAlaSerArgMetGluAsnGluArgThrLysAsp 511
QY 661 ATCTGTGTCACACCCAGTGTGTGTCATCATTCGAAAGAGGCTTGACCAATTAAG 720
Db 512 LeuIleIleGlnGluArgPheHisArgThrIleIleGlyGlnLysGlyGluArgIleArg 531
QY 721 AACATCATTAACAGACACCGATCCCGGAGTAAATCATCAAGAAAGAACTCTGAGCT 780
Db 532 GluIle--ArgAspLysPheProGluValIleIleAsnPheProAsp-----ProAla 548
QY 781 GCAGAGAACCTGTGCACCATCCATGCAACCCAGAGGAGACTTCTGAAGATGCCGATG 840
Db 841 ArgLysSerAspIleValGlnLeuArgGlyProLysAsnGluValGluLysCysThrLys 868
QY 841 ATCTTGAATCATGACAGAAAGGACGATGAGACCAAACTAGCCGAGAGATTCCTCTG 900
Db 569 TyrMetGlnLysMetValAlaAspLeuValGluAsnSerTyrSerIleSerValProIle 588
QY 901 AAAATCTTGACACACATGGCTGTGTGAAGACTGATGGAAAGAGCAGAAATTG 960
Db 589 PheLysGlnPheHisLys-----AsnIleIleLysGlyGlyAlaAsnIle 604
QY 961 AAGAAATTAACATGAAACAGGACCAAGCAAGTAAATCATCTTATCTTGACAGATTGAGC 1020
Db 605 LysLysIleArgGluLeuSerAsnThrLysIle-----AspLeuPro 618
QY 1021 ATATACACCCG--GAAAGACCATCATCTGTAAGGC-----ACAGTTGAGGCC 1068
Db 619 AlaGluAsnSerAsnSerGluThrIleIleIleThrGlyLysArgAlaAsnCysGluAla 638
QY 1069 TGTGCCAGTGCCTGATAGATTAAGAGAAAGCTGTGAGGCTTTGAAATGATATG 1128
Db 639 AlaArgSerArgIleLeuSerIleGlnLysAspLeuAlaAsnIleAlaGlu----- 655
QY 1129 CTGGCTGTAAACACCACTCCGATACTTCTTCAGACCTGTACCCCAACCAAGTTTGGC 1188
Db 655 ----- 655
QY 1189 CCGTCCCGCATCATCATCTTATCCAGAGCAGAGATGTGAATCTTCAATCCCAACC 1248

```

```
Db      656 -----ValGluValserIleProAla 662
QY      1249 CAGGCTGTGGGCGCATCATCGGGAAGAGGGGCACACATCAACAGCTGGCGAGATTG 1308
Db      663 LysLeuHisAsnSerLeuIleGlyThrLysGlyArgLeuIleArgSerIleMetGluGlu 682
QY      1309 GCCGAGGCTCTATCAAGATTGCCCCGTGGGAAAGGCCAGAGCTCAGCGAAMAGATGTG 1368
Db      683 CysGlyGlyValHisIleHisPheProValGluGlySer-----GlySerAspThrVal 700
QY      1369 ATCATCACCGGGCCACCGAAGCCCATGTTCAAGGCCAGGAGGATCTTTGGAAACTG 1428
Db      701 ValIleArgGlyProSerSerAspValGluLysAlaLysGlnLeu--HisLeu 719
QY      1429 AAAGAGAAACTTCTTTAACCCCAAGAGAAGTGAAGCTGGAAGCCCATATCAGAGTG 1488
Db      720 AlaGluGlu-----LysGlnThrLysSerPheThrValAspIleArgAla 734
QY      1489 CCTCTTTCACAGCTGGCCGGGTGATGTGCAAAAGGTGGCAAGCCGTGAACGAACTGCAG 1548
Db      735 LysProGluTyrHisLysPheLeuIleGlyLysGlyGlyLysIleArgLysValArg 754
QY      1549 AACTTAACCAAGTCGAGAAGTCATCGTCCTCGTGACCAACCGCAGATGAAATGAGGAA 1608
Db      755 AspSerThrGlyAlaArgValIlePheProAlaAlaGluAspLysAspGlnAsp----- 772
QY      1609 GTGATCGTCAGAATTATGGGCACCTTTTCTAGCCAGACTGCACAGCGCAAGATCAGG 1668
Db      773 --LeuIleThrIleIleGlyLysGluAspAlaValArgGluAlaGlnLysGlnLeuGlu 791
QY      1669 GAATTGTACCAACAGTG 1686
Db      792 AlaLeuIleGlnAsnLeu 797
```

Search completed: July 23, 2004, 11:18:18
Job time : 80.5997 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:51:55 ; Search time 28.8809 Seconds
(without alignments)
11838.037 Million cell updates/sec

Title: US-09-270-437D-8

Perfect score: 5886

Sequence: 1 ggcacgagagagagcagaga.....aaccttgtaaatgtttatttc 3283

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 11681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+n2p_model -DEV=xlp
-O/cmg2_1/USPPO.spool.p/US9270437/runat.23072004.095739.27413/app.query.fasta_1.5582
-DB=SwissProt.42 -QFMT=fastan -SUFFIX=isp -MINMATCH=0.1 -IOOPL=0 -IOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09270437 @CGN 1.1 57 @runat.23072004.095739.27413 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	279	4.7	572	1	FUB3_HUMAN
2	277.5	4.7	721	1	FUB2_RAT
3	261	4.4	643	1	FUB1_HUMAN
4	259.5	4.4	651	1	FUB1_MOUSE
5	259	4.4	707	1	FUB2_MOUSE
6	259	4.4	769	1	ZBP2_CHICK
7	241.5	4.1	606	1	Y475_ARATH
8	240	4.1	339	1	PCB3_MOUSE
9	234	4.0	339	1	PCB3_HUMAN
10	229.5	3.9	403	1	PCB4_HUMAN
11	229.5	3.9	413	1	PBP2_YEAST
12	225	3.8	356	1	PCB1_HUMAN
13	225	3.8	356	1	PCB1_MOUSE
14	223	3.8	1268	1	VGLN_HUMAN
15	219.5	3.7	474	1	NOA1_RAT
16	219	3.7	403	1	PCB4_MOUSE
17	213	3.6	492	1	NOA2_HUMAN
18	211.5	3.6	362	1	PCB2_MOUSE

19	210	3.6	365	1	PCB2_HUMAN	Q15366	homo sapien
20	207.5	3.5	493	1	NOA1_MOUSE	Q91xne	mus musculus
21	207	3.5	510	1	NOA1_HUMAN	P51513	homo sapien
22	207	3.5	1270	1	VLIN_CHICK	P81021	gallus gall
23	187.5	3.2	606	1	TDRH_HUMAN	O9y2w6	homo sapien
24	184	3.1	629	1	PAB2_ARATH	P47371	arabidopsis
25	180	3.1	1222	1	S160_YEAST	P06105	saccharomyc
26	173	2.9	463	1	ROK_HUMAN	Q07244	homo sapien
27	173	2.9	463	1	ROK_RABIT	O19049	oryctolagus
28	172	2.9	660	1	YH11_EBV	P03181	epstein-bar
29	170.5	2.9	381	1	YBD2_YEAST	P38199	saccharomyc
30	167	2.8	5147	1	PCLO_HUMAN	O9y6v0	homo sapien
31	166	2.8	2842	1	APC_RAT	P70478	rattus norv
32	160.5	2.7	470	1	NR54_HUMAN	Q15233	homo sapien
33	160.5	2.7	1781	1	AKI2_HUMAN	Q02952	homo sapien
34	157	2.7	1850	1	VIT2_CHICK	O02952	gallus gall
35	156	2.6	1943	1	PC15_MOUSE	O09p11	mus musculu
36	153.5	2.6	5262	1	MLT2_HUMAN	O14686	homo sapien
37	152.5	2.6	653	1	PABP_SCHPO	P31209	schizosach
38	152	2.6	892	1	ATX7_HUMAN	O15265	homo sapien
39	150.5	2.6	572	1	Y062_DEIRA	O9yrm6	deinococcus
40	149.5	2.5	867	1	ATX7_MOUSE	O8r411	mus musculu
41	149.5	2.5	3664	1	ELV3_HUMAN	O14576	homo sapien
42	148.5	2.5	367	1	MYT3_HUMAN	Q06158	homo sapien
43	148.5	2.5	511	1	YGB7_BACTN	O14576	homo sapien
44	148	2.5	1080	1	MYT3_MOUSE	O8y290	bacteroides
45	147	2.5	784	1	YAV2_XANCV	P59759	mus musculu

ALIGNMENTS

```
RESULT 1
FUB3_HUMAN
AC 096124; 092346; 095VB6; PRT; 572 AA.
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Far upstream element-binding protein 3 (FUSE binding protein 3).
GN FUBP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=97094955; PubMed=8940189;
RX Davis-Smyth T., Duncan R.C., Zheng T., Michelotti G., Levens D.;
RT "The far upstream element-binding proteins comprise an ancient family
RT of single-strand DNA-binding transcription activators.";
RL J. Biol. Chem. 271:31679-31687(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX TISSUE=Cervix, and Placenta;
RC MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uebis T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
```

human and mouse sequences." RT
Proc Natl Acad Sci U S A. 99:16899-16903(2002).
-1 FUNCTION: May interact with single-stranded DNA from the far-
CC upstream element (FUSE). May activate gene expression.
CC -1 SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1 ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=G96124-1; Sequence=Displayed;
CC Name=2;
CC IsoId=G96124-2; Sequence=VSP_008323, VSP_008324;
CC Note=No experimental confirmation available[1] lines.
CC -1 TISSUE SPECIFICITY: Detected in a number of cell lines.
CC -1 SIMILARITY: Contains 4 KH domains.
CC -1 CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 18.
CC -1 CAUTION: Ref.2 (AAH01325) sequence differs from that shown due to
CC a frameshift in position 527.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb-sdb.ch).
CC
DR EMBL; U69127; AAC50893.1; ALT FRAME.
DR EMBL; BC001325; AAH01325.1; ALT_FRAME.
DR EMBL; BC007874; AAH07874.1; --
DR Genew; HGNC:4005; FUBP3.
DR MIM; 603536; --
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 3.
DR SMART; SMO0322; KH; 3.
DR PROSITE; PS50064; KH TYPE 1; 4.
DR Transcription regulation; Trans-acting factor; Nuclear protein;
KW DNA-binding; Repeat; Alternative splicing.
FT DOMAIN 77 141 KH 1.
FT FT 162 228 KH 2.
FT FT 253 317 KH 3.
FT FT 354 421 KH 4.
FT FT VARSPLIC 1 64
MAEIVGASAPYGVGMKAEGFVDALHRVROLIAKIDISPHLN
SPLPYDPSYGVGVQKRPLDDGV -> MPI (in
isoform 2).
FT FT /FTid-VSP_008323.
FT FT Missing (in isoform 2).
FT FT /FTid-VSP_008324.
FT FT V -> D (IN RBF. 1).
FT FT CONFLICT 72 72
SQ SEQUENCE 572 AA; 61640 MW; FIBB225542BC197D CRC64;

Alignment Scores:
Pred. No.:
Score: 2.13e-10 Length: 572
Percent Similarity: 279.00 Matches: 117
Conservative: 39.96% Conserved: 82
Best Local Similarity: 23.49% Mismatches: 202
Query Match: 4.74% Indels: 98
DB: 1 Gaps: 16

US-09-270-437D-8 (1-3283) x PUB3_HUMAN (1-572)
QY CATCACTGGAGAACACTACCTTCAGATTTCCTCATCCCGGANTGAAGGTAGAGCTCC 555
Db |||::: ::||| |
24 HHSARYVALARGNLNLEALAALAYSLIAPSSEILLEPRO-----HLSLEAADAAN 41
QY CCTTGCCCCCTCAGGAGCCGCGCTG-----GGACAACCTCTTC 597
Db :::| | | | | :|||
42 SetThr-ProLeuValAspProSerValTYRGIYTRYGLYVALGNLYSARGPROLEUAS 61
QY CGGAGACAAAGGCCAGCCCTCGGGGGCAGCTTCACGAGCAGAACAATGATTTCCGGCTG 657
Db 598 CGGAGACAAAGGCCAGCCCTCGGGGGCAGCTTCACGAGCAGAACAATGATTTCCGGCTG 657
61 pAepgIyAlglYAmgnInleugIyAlaeuValIHsglnArgThrvAllelrnglu-- 80

QY	658	CGAGTCTGTCCTCCCAACCCAGTTGTTGGTCATCATCGAAGAGGGCTTGACCAAT	717
Db	81	GluphelysValProspysMetValGlyPheIleIleGlyArgGlyGlyValGluIle	100
QY	718	AAGAACTTCACTAAGAGAGACCCAGTCCCGGATGATTCATCATAGAAAGAACTCTGGA	777
Db	101	SerArgIleGlnIleGluSerGlyCysIysIleGlnIle--AlaSerGluSerSerGly	119
QY	778	GCTGCAGAGAGCGCTGTCCATCATTCATGCCACCCCAAGAGGAGACTTTCAGACATGCCG	837
Db	120	IleProGluArgProCysValIleThrGlyThrProGluSerIleGluGlnAlaIysArg	139
QY	838	ATAGTCTTCAAAATTCATGACAGAAA-----GAGCAGAT	870
Db	140	LeuIeuGlyIleIleValaIaPargCysArgAsnGlyProGlyPheHisAsnSerIleasp	159
QY	871	GAGACCAACTAGCCGAGAGATTCCTCTGAAAATCTTGCGACACATAGGCTTGTTTGA	930
Db	160	SerAsnSerThrIleGlnGlu-----IleIeuIleProAlaSerIysValGly	175
QY	931	AGACTGATTTGAAAAGAGAGACAGAAATTGAAAGAAATGACATGAAACAGAGACCAAG	990
Db	176	LeuValIleGlyArgGlyGlyGluThrIleLysGlnIleGlnGluArgThrGlyValIys	195
QY	991	ATAACAATCTCATCTTTTGACAGATTGAGCATA--TACAAACCGGAAAGAACCATCACT	1047
Db	196	MetValMet-----IleGlnAspGlyProLeuProThrGlyIaIaSpIysProLeuArg	213
QY	1048	GTGAAGGCGACAGTTGAGGCGCTGTGCCAGTCGAGATGAGATGAGATGAAAGACATCCGT	1107
Db	214	IleThrGlyAspAlaPheLysValGlnGlnAlaArgGluMetValIleuGluIleIleArg	233
QY	1108	GAGGCTTTGAAATGATATGCTGGCTGT-----AACACCACTCCGGATAC	1155
Db	234	GluIysAspGlnAlaAspPheArgGlyValArgGlyAspPheAsnSerArgMetGlyGly	253
QY	1156	TTTTCAGCCTGTGACCCCATCACCAAGTTGGCCCGTCCCGCATCACTCTTANCCA	1215
Db	254	GlySer-----	255
QY	1216	GAGCAGAGAGATTGGAAATCTTTCATCCCAACCGAGCTGTGGCGCATCATCGGAG	1275
Db	256	-----IleGluValSerValProArgPheAlaValGlyIleValIleGlyArg	271
QY	1276	AAGGGGACACATCAAAACAGCTGGAGATTCGCCGAGACCTCTATCAAGATTGCCCT	1335
Db	272	AsnGlyIleMetIleLysIleIleGlnAspAlaGlyValArgIleGlnPheLysPro	291
QY	1336	GCGAGAGGCCAGACAGTCAGAGAAAGATGTCATCAACGGGCGACCGGAGGCCAG	1395
Db	292	AspAspGlyIleSer--ProGluArgAlaAlaGlnValMetGlyProProAsp--Arg	309
QY	1396	TTCAAGGGCCACGAGCGAGATCTTTGGAAACTG-----AAAGAGAAAACCTTC	1443
Db	310	CysGlnHisAlaIleHisIleIleSerGluIleuIleLeuThrIleGlnGluArgAspGly	329
QY	1444	TTTAAACCCAAAGAAAGAGTAG-----	1467
Db	330	PheGlyGlyLeuAlaIleAlaIleArgGlyIleArgGlyIleArgGlyIleAspTyrSerValGly	349
QY	1468	-----CTGGAAGCGATTCAGAGTGCCTCTTCCACACAGTGCACGCGGTG	1512
Db	350	AlaProGlyGlyValGlnGluIleThrIleThrValProAlaAspLysCysGlyIleuVal	369
QY	1513	ATTGGCAAAAGTGGCAAGACCGTGAACGACTGCAGAACTTAACCAAGTCGAAAGTCACT	1572
Db	370	IleGlyLysGlyGlyIleAsnIleLysSerIleAsnGlnIleSerGlyIleHisIleValGlu	389
QY	1573	GTGCTCTGTGACCAACCGCAAGTAAATGAGAAAGATGTCAGAAATTATCGGGAC	1632
Db	390	LeuGlnArgAsnProProAsnSerAspProAsnIle-----	402

```

QY 1633 TTCTTTGTAAGCCAGCTGCACAGCGCAAGATCAGGGAATTGTACAGAGTGAACGACG 1692
DB 403 -----ArgArhPheThrIleArgGlyValProGlnGlnIleGluVal 416
QY 1693 CAGGAGCAGAAATATCCC-----TAGGAGTGCCTTCACAGCGCAG----- 1733
DB 417 A1aArgGlnLeuIleAspGlnbysValGlyGlyThrAsnLeuGlyAlaProGlyAlaPhe 436
QY 1734 ---CAAGTAGAGCTCCCAAGGACGACGCAAAACAAGATGAATGTAGCCCTTCACACA 1790
DB 437 GlyGlnSerProPheSerGlnProProAlaProProHisGlnAsnThrPheProProAla 456
QY 1791 CCTGACAGAAATGACCAAAACGACGACCAAGCTGAGGAGCAACCA 1838
DB 457 SerSerGlyCysPheProAsnMetAlaAlaLysValAsnGlyAsnPro 472

RESULT 2
FUB2_RAT
ID_FUB2_RAT STANDARD; PRT; 721 AA.
AC Q99PF5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Far upstream element binding protein 2 (FUSE binding protein 2) (KH
DE type splicing regulatory protein) (KSRP) (MAP2 RNA trans-acting
DE protein 1) (MARTAL).
GN FUBP2 OR KHSRP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 73-87; 89-108 AND 475-486, SUBCELLULAR
RP LOCATION, AND FUNCTION.
RC TISSUE-Brain.
RX MEDLINE=22246918; PubMed=12358751;
RA Redben M., Wege K., Buck F., Schweizer M., Richter D., Kindler S.;
RT "Molecular characterization of MARTAL, a protein interacting with the
RT dendritic targeting element of MAP2 mRNAs."
RT J. Neurochem. 82:1039-1046(2002).
CC -!- FUNCTION: Part of a ternary complex that binds to the downstream
CC control sequence (DCS) of the pre-mRNA. Mediates exon inclusion in
CC transcripts that are subject to tissue-specific alternative
CC splicing. May interact with single-stranded DNA from the far-
CC upstream element (FUSE). May activate gene expression (by
CC similarity). Binds to the dendritic targeting element and may play
CC a role in mRNA trafficking.
CC -!- SUBUNIT: Part of a ternary complex containing FUBP2, PTBP1, PTBP2
CC and HNRPH1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear. A small proportion is also found in
CC the cytoplasm of neuronal cell bodies and dendrites.
CC -!- SIMILARITY: Contains 4 KH domains.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF308818; AAG59811.1; -
DR InterPro: IPR004087; KH dom.
DR InterPro: IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 4.
DR SMART; SM00322; KH; 4.
DR PROSITE; PS50084; KH_type_1; 4.
KW Transport; mRNA transport; mRNA processing; mRNA splicing;
KW Transcription regulation; Trans-acting factor; Nuclear protein;
KW DNA-binding; RNA-binding; Repeat.
FT DOMAIN 145 209 KH 1.
FT DOMAIN 234 300 KH 2.
FT DOMAIN 323 387 KH 3.

```

```

FT DOMAIN 425 492 KH 4.
FT REPEAT 572 685 4 X 12 AA IMPERFECT REPEATS.
FT REPEAT 572 583 1.
FT REPEAT 618 629 2.
FT REPEAT 644 655 3.
FT REPEAT 674 685 4.
FT DOMAIN 7 68 GUY/PRO-RICH.
FT DOMAIN 69 498 GUY-RICH.
FT DOMAIN 499 613 ALA/GUY/PRO-RICH.
SQ SEQUENCE 721 AA; 74226 MW; 482C7A765C60EE4A CRC64;

Alignment Scores:
Pred. No.: 2,84e-10 Length: 721
Score: 277.50 Matches: 161
Percent Similarity: 34.08% Conservative: 81
Best Local Similarity: 22.68% Mismatches: 283
Query Match: 4.71% Indels: 187
DB: 1 Gaps: 27

US-09-270-437d-8 (1-3283) x FUB2_RAT (1-721)
QY 104 CGCGCGTCACGCGCGAGCAGCTCCGCGCAGCTCTTGGGACAGAGAGCTGCCGCG 163
DB 10 ProProGlyProProProProProAlaGlyGlyGly-GlyGlyAlaAlaGlyAlaGlyGly 29
QY 164 GACAGGTCTGCTGAAGTCCGCGCTACGCTTGAGACTACCCGACGACAACTGGGCA 223
DB 29 yProProGlyProProGlyAlaGlyAlaGlyAspArgGlyGlyGlyProGlyGlyGly 49
QY 224 TCCGGGCATGAGACGCTCTCGGGTAAAGTGAAATTCATGGGAAATCATGAAAGTTG 283
DB 49 yPro-----GlyGlyGlyGlyAla----- 55
QY 284 ATTACTAGTCTTAAAGGCTAAAGAGACAGAAATTCAGATTGAAACATCCCTCTC 343
DB 56 -----SerGlyGlyProSerGly 61
QY 344 ACCTCAGGTGGAGGTGTTGATGACTTTG-----GCTCAATATGAGACAGTGGAGA 397
DB 61 nProProGlyGlyGlyGly-ProGlyIleArgLysAspAlaPheAlaAspAlaValGln 81
QY 398 ATGTGAACA-----GTCACACAGACACAGAAACCGCGTTGTCAAGTCA 445
DB 81 rgAlaArgGlnIleAlaAlaLysIleGlyLysAlaAlaIleThrValAsnAsn 101
QY 446 CATATGCAACAAGAGAAGAACAAATATAGCATGAGAGCTAAGCGGATCAGTTG 505
DB 101 hr-----P 102
QY 506 AGAATCACTCTTCAAGATTCTTCAATCCCGATGAGAAGTGAGTCCCTTCGCCCC 565
DB 102 roAspPheGlyPheGlyGlyGlnLysArgGlnLeuGlnAspGlyAspGlnProAspSer 122
QY 566 CTGACGAGCGCGAGCGGGGACCATCTTCCGGGAGCAAGC-----CAGCGCTGTG 619
DB 122 ysisIleuAlaSerGlnGlyAspSerIleGlySerGlnLeuGlyProIleHisProPro 142
QY 620 GGGGCACTTCTCAAGCGCAGACAGATTGATTTCCGCTGCGGATCTCTGTCGCCACCCAGT 679
DB 142 roArhThrSerMetThrGlnGlu-----TyrArgValProAspGly 156
QY 680 TTGTTGGTGCATCATCGAAGAGGCTTGACCATTAAGAACATCATCATGACGACGCC 739
DB 156 etValGlyLeuIleIleGlyArgGlyGlyGlnIleAsnLysIleGlnLysPerg 176
QY 740 AGTCCGGGTGAATTCATAGAAAGAAAGAACTGTGAGCTGACAGAGAAGCTTCAACA 799
DB 176 lyCysLysValGlnIle---SerProAspSerGlyGlyLeuProGlnLysSerValSer 195
QY 800 TCCATGCAACCCGACGAGGAGCTTCTGAACATGACGATGATCTTGAATCATCATGAGA 859
DB 195 eurThrGlyAlaProGlnSerValGlnLysAlaLysMetLeuAspAspIleValSer 215

```


RT "Identification of human DNA helicase V with the far upstream
RT element-binding protein.";
RL Nucleic Acids Res. 29:1061-1067 (2001).
RN [4]
RP PARTIAL SEQUENCE, AND MASS SPECTROSCOPY.
RX MEDLINE=22166132; PubMed=12176931;
RA Rapsilber J., Ryder U., Lamond A.I., Mann M.;
RT "Large-scale proteomic analysis of the human spliceosome.";
RL Genome Res. 12:1231-1245 (2002).
RN [5]
RP INTERACTION WITH SIABP1.
RX MEDLINE=20337922; PubMed=10882074;
RA Liu J., He L., Collins I., Ge H., Libutti D., Li J., Egly J.-M.,
RT "The FBP interacting repressor targets TFIIH to inhibit activated
RT transcription.";
RL Mol. Cell 5:331-341 (2000).
RN [6]
RP INTERACTION WITH JTV1, UBIQUITINATION, AND PROTEASOME-MEDIATED
RP DEGRADATION.
RX MEDLINE=22716800; PubMed=12819782;
RA Kim M.J., Park B.-U., Kang Y.-S., Kim H.-J., Park J.-H., Kang J.-W.,
RT Lee S.-W., Han J.-M., Lee H.-W., Kim S.;
RT "Downregulation of FUSE-binding protein and c-myc by RNA synthetase
RT cofactor p38 is required for lung cell differentiation.";
RL Nat. Genet. 34:330-336 (2003).
RN [7]
RP STRUCTURE BY NMR OF 278-447 IN COMPLEX WITH SINGLE STRANDED FUSE DNA.
RX MEDLINE=21864720; PubMed=11875576;
RA Braddock D.T., Louis J.M., Baber J.L., Levens D., Clore G.M.;
RT "Structure and dynamics of KH domains from FBP bound to
RT single-stranded DNA.";
RL Nature 415:1051-1056 (2002).
CC -1- FUNCTION: Regulates MYC expression by binding to a single-stranded
CC far-upstream element (FUSE) upstream of the MYC promoter. May act
CC both as activator and repressor of transcription.
CC -1- SUBUNIT: Interacts with SIABP1/FIR and JTV1.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96AE4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q96AE4-2; Sequence=VSP_008321;
CC Note=No experimental confirmation available;
CC -1- PTM: Ubiquitinated. This targets the protein for proteasome-
CC mediated degradation.
CC -1- SIMILARITY: Contains 4 KH domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U05040; AAA17976.2; -;
DR EMBL: BC017247; AAA17247.1; -;
DR PDB: 1J4W; 06-MAR-02.
DR Genew; HGNC:4004; FUBP1.
DR MIM: 603444; -;
DR InterPro: IPR008160; Collagen.
DR InterPro: IPR004087; KH dom.
DR InterPro: IPR004088; KH_Type_1.
DR Pfam: PF00013; KH; 4.
DR SMART: SM00322; KH; 4.
DR SMART: SM00322; KH; 4.
DR PROSITE: PS50084; KH_Type_1; 4.
KM Transcription regulation; Trans-acting factor; Nuclear protein;
KM DNA-binding; Repeat; Ubl conjugation; Alternative splicing;
FT 3D-structure. 99 163 KH 1.
FT DOMAIN 184 250 KH 2.

FT DOMAIN 274 338 KH 3.
FT DOMAIN 375 442 KH 4.
FT DOMAIN 13 26 GLY-RICH.
FT DOMAIN 348 395 GLY-RICH.
FT DOMAIN 449 559 PRO-RICH.
FT VASPLIC 642 643 CQ -> CREPDASTIAL (in isoform 2).
FT CONFLICT 96 96 /FTID=VSP_008321.
SQ SEQUENCE 643 AA; 67473 MW; 086DAEAA0ACF807B CRC64;
Alignment Scores:
Pred. No.: 3.2e-09 Length: 643
Score: 261.00 Matches: 140
Percent Similarity: 37.56% Conservative: 85
Best Local Similarity: 23.37% Mismatches: 237
Query Match: 4.43% Indels: 137
DB: 1 Gaps: 25
US-09-270-437D-8 (1-3283) x FUBP_HUMAN (1-643)
QY 336 CCTCTCCTCAGCTGAGAGTGTGGATGACTTTGGCTCAATATGGACAGTGA 395
DB 10 ProterseergerlyserAlaGlyGlyGly-----GlyGlyGly 22
QY 396 GAATGTGAACAGTCAACACACACACACAAACCGCGTGTCAACGTCACTATGCAAC 455
DB 23 GlyGlyGly-GlyValAsnAspAlaPheIysAspAlaLeuGlnArg-----AlaAr 39
QY 456 AAGGAAGAAGCAAAATACCATCGAGAGCTAAGCGGGCATCAGTTTGAAACTACTC 515
DB 39 gglnlleAlaAlaIysIleGlyAspAlaGlyThSerleuAsnSerAsnAspIYrG1 59
QY 516 CTTCAAGATTTCATCATCCCGATGAAGAGTGAAGTCCCTCGCCCGCCAGCAGGC 575
DB 59 YTYrGlyGlyGlnIlyAsrProleuGlnAspGlyAspGlnProAspAlaIylsValAl 79
QY 576 CCAAGCGGGGACCATCTTCCCGGAGCAAGCCACCCCTGGGGGCACTTCTCAGGC 635
DB 79 aProGlnAsnAspSerPheGlyThrGln-----LeuProPromeThIsgInGlnG1 96
QY 636 CAGACAGATTGATTCCTCCGCTGGGAGATCCGTGCCACCCAGCCAGTTTGGTCATCAT 695
DB 96 nArGserVal---MetThrIgluIuTYrIysValProAspGlyMetValGlyPheIleI1 115
QY 696 CGGAAAGAGGGCTTGAACATTAAGAAACATCACTAAGACAGACCATCGCGGATGAT 755
DB 115 eGlyArGlyGlyGlnGlnIleSerArGlnIleGlnGlnIleSerGlyIylsIleGlnI1 135
QY 756 CCATGAAGAAAGAACTCTGAGAGCTGACAGAAAGCTGTCAACATCCATCCACCCAGA 815
DB 135 e--AlaProAspSerGlyIleuProGlnArgSerCyMetLeuThrGlyThrProG1 154
QY 816 GGGGACTTCTGAAGCATGCGCATGATTTCTTGAATCATGCAAAA-----861
DB 154 uSerValGlnSerAlaIlyAsrIleuAsnGlnIleValGlnIylsIylsValrProAlaPr 174
QY 862 -----GAGCAGATGAGACCAACTAGCGGAAGATTCTCTGAAAACTTGGC 911
DB 174 oGlyPheHishIsgIylAspGlyProGlyAsnAlaValGln-----GluIleMetI1 191
QY 912 ACACAAATGGCTGTGTTGGAAGTCACTGTAAGAAAGAGCAGAAATTTGAAGAAATGA 971
DB 191 eProAlaSerIylsAlaGlyIleuValIleGlyIylsGlyGlyGlnThrIleIylsGlnIeug1 211
QY 972 ACATGAAGACAGGACCAAGATTAACAATCATCTTTGAGAGATTGACATATACAAACC 1031
DB 211 nGluArGlnAlaGlyValIylsMetValMet-----IleGlnAsp-----GlyPr 225
QY 1032 GAAAAAGAAC-----ATCACTGTGAAGGCGACAGATTGAGGCTGTGCAG 1076
DB 225 oGlnAsnThrGlyAlaAspIylsProleuArGlnIleThrGlyAspProTYrIylsValGlnG1 245
QY 1077 TGCTGAGATGAGATTTATGAGAAAGCTGCGTGTGAGGCTTTGAAATGATATGCTGCTGT 1136

```

Db      245  |||...  :::::  ::|||...  -----  256
           |||...  ::|||...  -----  256
           |||...  ::|||...  -----  256
Db      1137  TAAGACCCACATCCGATACCTTCCAGCTTACCCCATCAGATTGGCCGCTTCC  1196
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      257  -----  GlnlyglPheAagGluVal-----  ArgsmgluYrgly-----  268
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1197  GCATCATCATCTTATTCAGAGCAGAGATTGTGATCTCTTCACTCCACCCAGGCTGT  1256
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      269  -----  SerArglleglYglYasnglYlleasprAlProleProAaPheAlaVa  286
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1257  GGGGGCCATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1316
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      286  lglYlleValleaglYasnglYglumetlleYsYsllleGlnasnaPheAlaGlyVa  306
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1317  CTCTATCAAGATTGGCCCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  1376
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      306  lArglleGlnPheYsProAasprGly-----  ThrThrProGluYrglleAlaGlnlleTh  325
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1377  CGGGCCACCGGAA-----  GCCCAGTTCAAGGCCAG-----  1407
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      325  rglYProProAasprArgYsGlnHleAlaGlnllelleThraPheLuleuYsSerVa  345
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1408  -----  GGAGCGATCTTTGGGAAA-----  1425
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      345  lGlnAlaGlyasnProGlyglYProGlylYProGlylYArgGlylYArgGlylYlGlnl  365
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1426  -----  CTGAAGAAGAAAATTCTTTAAACCCCAAGA  1457
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      365  YAsnTrpAsnMetGlyProProGlylYleuGlnGlnPheAsnPhelle-----  381
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1458  AGAAGTGAAGCTGGAAGGCGATATCAAGATGCCCTTTCCAAGCTGGCGGGTATTGG  1517
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      382  -----  ValProThrGlylYsThrGlyleuIlelel  392
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1518  CAAAGGTGCAAGACCGTGAACGAATCTGAAGAACTTAACCAAGTGAAGTGTGCGC  1577
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      392  YlYsGlyglYglumetlleYsSerlleSerGlnGlnInsrglylAlaArglleGlnleu  412
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1578  TCGTGAACCAAGCGCAGATGAATGAAGGAAGT-----  ATGTCAGAAATTATGGGCAC--  1632
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      412  nArgsnProProProAasnaAlaAprAasMetYsleuPheThlleArgGlylYhrPr  432
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1633  -----  TTCTTTGTACGACGATCTGCACAGCGCAAGATCAAGAGGAATTGTACACA  1682
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      432  oGlnGlnlleasprYlAlaArglleuIleleuGlnYsllleGlyglYProValasnPr  452
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      452  oleuGlyProProValProHleGlylYProHleGlylYProGlylYProProG  472
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1743  CTCCACAGAGCAGCAGCAAAAACAGAGATGTATGCCCTTCAACACTCTGACAGAAATG  1802
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      472  lYProProGly-ProGly-----  ThrPro-----  Met  480
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1803  AGACCAAGCAGCAGCAGCAGATCGGAGAGCAACCAAGACCATGTAGAGATGAGAAGT  1862
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      481  GlYProYlYasnProAlaProYlYasnProGlyPro-----  492
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1863  CTGCGAGGCGCGCAGACTTGCAGAG-----  CCCTGAGAACCCCAAGGGCCGAGAGAG  1919
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      493  -----  ProGlyProAlaProHleGlylYProProAlaProYlYalProGlnGly  508
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      1920  GCGCGGAGAGGTAGCCAGGTTGGCCAGAA-----  CCACCGAGCCCGCGC  1964
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256
Db      509  TrpGlyAsnaAlaYlYProHleTrpGlnGlnAlaAlaProProAasProAla  525
           |||...  |||...  ::|||...  -----  256
           |||...  |||...  ::|||...  -----  256

```

```

DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Far upstream element binding protein 1 (FUSE binding protein 1) (FBP).
GN      FUBP1 OR D3ERTD330E.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
[1]
RN      SEQUENCE FROM N.A. (ISOFORM 1).
RP      TISSUE=Retina.
RC      MEDLINE=22388257; PubMed=12477932;
RX      Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner U., Sherman C.M., Schuller G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Ditchenko L., Marusina B., Farmer A.A., Rubin G.M., Hong L.,
RA      Stempleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Prange C.,
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Schechenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA      Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RN      SEQUENCE OF 1-286 FROM N.A. (ISOFORM 2).
RP      STRAIN=C57BL/6J; TISSUE=Head;
RX      MEDLINE=22354683; PubMed=12466851;
RA      Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA      Nikaido I., Osato N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,
RA      Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach G., Gotohori T.,
RA      Baldarelli R., Hill D.P., Bull C., Hume D.A., Quackenbush J.,
RA      Schiml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA      Blake J.A., Brad D., Busic V., Chochia C., Corbali L.E., Cousins S.,
RA      Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA      Gaasterland T., Gariboldi M., Gissi C., Godik A., Gough J.,
RA      Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA      Kanai A., Kawaji H., Kawasawa Y., Kedzielski R.M., King B.L.,
RA      Kongave A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A.,
RA      Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA      Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA      Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA      Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA      Sadelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA      Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA      Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA      Wilming L.G., Wyszewski-Boris A., Yanagisawa M., Yang L., Yang L.,
RA      Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA      Hirovane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA      Shiraki T., Waki K., Kawai J., Aizawa K., Aizawa T., Fukuda S.,
RA      Hara A., Hashizume W., Imotani K., Ishii Y., Itoh N., Kagawa I.,
RA      Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA      Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA      Birney E., Hayashizaki Y.;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573 (2002).
RC      -1- FUNCTION: Regulates MYC expression by binding to a single-stranded
CC      far-upstream element (FUSE) upstream of the MYC promoter. May act
CC      both as activator and repressor of transcription (By similarity).
CC      -1- SUBUNIT: Interacts with SIN3BP1/FR and JTV1 (By similarity).
CC      -1- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
CC      Name=1;
CC      IsoId=Q91WU8-1; Sequence=displayed;

```


QY 14 GACAGTCTCTGTAAGTCGCGCTACGCTTCGTGAGTACCCTGAGAACTGGGCA 223
Db 29 yProProProGlyProProGlyAlaGlyAspArgGlyGlyGlyProCysGlyGly 49
QY 224 TCCGGCCATCGAACCCTCTCGGGTAAAGTGAATGCATGGGAAAAATCATGAAGTTG 283
Db 49 yPro----- 50
QY 284 ATTACTAGTCTCTAAAAAGCTAAGAGAGAAAATTCAATTGGAACATCCCTCTC 343
Db 51 -----GlyGlyGlySerAlaGlyGlyProSerG 60
QY 344 ACTCGACGTGGAGGTGTGATGAGCTTTG-----GCTCAATATGAGACAGTGA 397
Db 60 nProProGlyGlyGly- ProGlyIleArgGlyAspAlaPheAlaAspAlaValGln 80
QY 398 ATGAGAACTCAACACAGACACAGAAACCGCGTTGTCAAGTCAATATGACAA 457
Db 80 rgaIaArgGlnIle----- 84
QY 458 GAGAAAGCAAAAATAGCCATGAGAGCTAAGC-----GGCATCAGTTTGAAGACT 511
Db 85 -----AlaAlaAlaGlyIleGlyAspAlaAlaThrThrGlyAsnAsnSerThrProAsp 103
QY 512 ACTCTTCAGATTTCTACATCCCGATGAAGAGTGAAGTCCCTTCGCGCCCTCAGC 571
Db 103 heGlyPheGlyGlyGlnAspArgGlnLeuGlnAspGlyAspGlnProGlySerIleGly 123
QY 572 GAGCCCAAGCTGGGACCACTCTTCGCGAGCAAGC-----CACGCCCTGGGAGCA 625
Db 123 euAlaSerGlnGlyAspSerIleSerSerGlnLeuGlyProIleIleProProAspArg 143
QY 626 CTTCACAGGCGACAGATTTATTTCCCGCTCGGATCTGTCTCCCAACCCAGTTTGTG 685
Db 143 hrSerMetThrGlnIle-----TyrArgValProAspGlyMetValG 157
QY 686 GTGCATCATCGAAGAGAGGCTTGACCATTAAGAACATCATCAATAGAGACCAAGCCC 745
Db 157 IlyLeuIleIleGlyArgGlyGlyGlnGlnIleAsnIleGlnGlnAspSerIleCysL 177
QY 746 GGGTATATTCATTAAGAAAGAACTGTGAGCTGAGAGAAAGCTGTGACCATTCATG 805
Db 177 yValGlnIle-----SerProAspSerGlyGlyLeuProGlnArgSerValSerLeuArg 196
QY 806 CCACCCAGAGGGGACTTTCGAGCATGCCAGATGTTCTTGAATCATGCAGAAA---- 861
Db 196 IyAlaProGlnSerValGlnIlySalIlyMetMetLeuAspIleValSerArgGly 216
QY 862 -----GAGCAGATGAGACCAACTAGCCG 886
Db 216 rglGlyGlyProProGlyGlnPheHisAspAsnAlaAsnGlyGlyGlnAsnGlyThrValG 236
QY 887 AAGAGATTCCTCTGAAAAATTTGGCACACAATGGCTTGTGGAAGACTGATGGAAG 946
Db 236 ImlGlnIleMetIleProAla-----GlyLysAlaGlyLeuValIleGlyLysG 252
QY 947 AAGCGAATTTGAGAAAATTTGAACATGAAACAGGACCAAGATACATC-----T 1000
Db 252 IyGlyGlnThrIleLysGlnLeuGlnIlyArgAlaGlyValIlyMetIleLeuIleGln 272
QY 1001 CATCTTTGAGATTTGAGCATATACAAACCGGAAAACCATCAGTGTGAAGGACAG 1060
Db 272 spGlySerGlnAsnThrAsnValAspLysProLeuArgIleIleGlyAspProTyrLysV 292
QY 1061 TTGAGGCTGTGCAAGTGTGAGATGAGATTATGAGAGACTGCGTGAAGCTTTGAAA 1120
Db 292 alGlnGlnAlaCysGlnMetValMetAspIleLeuArgAsnValThrIlySalGlyPheG 312
QY 1121 ATGATATGCGGTGTTAACACCACTCCGAGTATCTTCACAGCTGTACCCCATCAGC 1180
Db 312 IyAsp-----ArgAsnG 316

QY 1181 AGTTTGC-----CCGTCCCGCATATCATCTTATC 1213
Db 316 IlyIyGlySerArgIleGlyGlyIleAspValProValProArgHis----- 332
QY 1214 CAGAGCAGAGATTTGAAATCTTCATCCCAACCAAGCTGTGGGCGGCATCGGGA 1273
Db 333 -----SerAlaGlyValValIleGly 340
QY 1274 AGAAGGGGACACATCAACAGCTGGAGATTCGCGGAGCTCTTATCAAGATTGCCC 1333
Db 340 rGserGlyGlnMetIleLysIleGlnAsnAspAlaGlyValArgIleGlnPheLysG 360
QY 1334 CTGGGAAAGGCCAGAGCTCAGCCGAAAGATGTCATATACCGGGCCACCGGAAGCCC 1393
Db 360 IAspAspGlyThrGly--ProGlnIyIleAlaHisIleMetIlyProProAsp---A 378
QY 1394 AGTTCAAGGCCACGAGCATCTTT----- 1419
Db 378 rGyGlnHisAlaAlaArgIleIleAsnAspLeuLeuGlnSerLeuArgSerGlyProP 398
QY 1420 -----GGGAACTGAAGAGAAACT 1441
Db 398 roGlyProProGlyGlyProGlyIleProGlyIyArgGlyArgGlyArgGlyGlnGly 418
QY 1442 TCTTAAACCCCAAGAAAGTGAAGCTGGAAGCCATATACAGTGCCTCTTCACAG 1501
Db 418 snIyPlyProGlyGlyGlnMetThrPheSer-----IleProIleHisLysC 434
QY 1502 CTGCGCGGTGATTTGCAAAAGTGAAGCTGGAAGCACTGCAGAACTTAACCACTG 1561
Db 434 yGlyGlnValIleGlyArgGlyGlyGlnValIlySalIleAsnGlnIleThrGly 454
QY 1562 CAGAACTCATGTCCTGTGACCAAGCCCA-----GATGAATAAGAGAAATGA 1612
Db 454 IAspValGlnIleSerArgGlnLeuProProIleGlyIleProIleSerIleuPheI 474
QY 1613 TCGTCAGA-----ATTACGGGACTCTTTGCTAGCCAGATGACAGCGCA 1660
Db 474 IleIleArgGlySerProGlnGlnIleAspHis-----CysArgGlnLeuIleGlnIle 492
QY 1661 AGATCAGGGAAT-----TGTAACAAGTGAAGCAGCAGAGAGCAAGAAATCCCTCAGGA 1716
Db 492 yIleGlnGlyProLeuCysProValGlyProGlyGlyProGlyProAlaGly 512
QY 1717 GTGCGCTCAGAGCGCAAGTAGAGCTCCACAGCAGCAAGAAACAGGATGAATG 1776
Db 512 roMetGlyProPheAsnProGlyProPheAsnGlnGlyProProGlyAlaProProHis 532
QY 1777 TAGCCCTTCACACCTGCAGCAATG----- 1802
Db 532 IagIyGlyProProIleHisGlnIlyProProGlnIlyTrGlyAsnThrIlyProGlnT 552
QY 1803 -----AGACCAAGCCAGCCAGCCAGATCGGAGCAACCAAGAA 1842
Db 552 IyGlnProProAlaProHisAspProSerIlySalAlaAlaAlaAlaAspProAsn 572
QY 1843 CCATCTGAGGA-----ATGAGAAGTCTGCGAGCGCGCCAGAGACTCTCCAGGC 1893
Db 572 IaAlaTrpAlaAlaIyTrIySerHisIyTrIyGlnIleProProGlyProValPro----- 590
QY 1894 CTTGAAGACCCCGAGGCGGAGAGGGCGGGAAGGTCAAGCGAGTTTCCAGAACAC 1953
Db 591 -----GlyProAlaProAlaProAlaProAlaIyGlnGlyProP 606
QY 1954 CGAGCCCGCTCCGCGCCCGCAG 1977
Db 606 roGln-ProProProThrGlyGln 613
RESULT 6
ZBP2_CHICK
ID ZBP2_CHICK STANDARD; PRT; 769 AA.
AC Q8UV09;
DT 26-FEB-2003 (Rel. 41, Created)

FT DOMAIN 138 210 KH 2.
 FT DOMAIN 311 380 KH 3.
 FT DOMAIN 394 455 KH 4.
 FT DOMAIN 535 599 KH 5.
 FT VARSPIC 532 532 L->F (in isoform 2).
 FT VARSPIC 532 532 /Frid=VSP_008899.
 FT VARSPIC 533 606 Missing (in isoform 2).
 FT VARSPIC 533 606 /Frid=VSP_008900.
 SQ SEQUENCE 606 AA; 65760 MW; 61F135BB8647C0C CRC64;
 Alignment Scores:
 Pred. No.: 5.71e-08 Length: 606
 Score: 241.50 Matches: 104
 Percent Similarity: 39.50% Conservative: 86
 Best Local Similarity: 21.62% Mismatches: 178
 Query Match: 4.10% Indels: 113
 Gaps: 16
 DB: 1
 US-09-270-437d-8 (1-3283) x Y475_ARATH (1-606)
 QY 208 GACCAAGAACTGAGCCATCCGCGCATGAGACCTCTCGGCTAAAGTGAATTGATGGG 267
 DB 136 AspspAenLysGluCyArgLeuValProPheSerGlnSerSerLeuLleGly 155
 QY 268 AAATCATGAGATTGATTACTAGCTCTAAAGCTTAAG3--AGCAGAAATTCAG 324
 DB 156 LysAlaGlyGlnAsnLleYsArgLleArgArgArgThrArgAlaSerValValVal 175
 QY 325 ATTCGACATC-----CCTCCATCCTGAGAGTGGAGGTGTGATGACTTTTGCT 378
 DB 176 SerLysAspValSerAspProSerHisValCysAlaMetGluTyrAspAsnValVal 195
 QY 379 CAATATGGAGAGTGAAGATGTGAACAAGTCAACACAGACAGAAACCCGCTTGC 438
 DB 196 IleSerGlyGlnProGlnSerValLysGln-----AlaLeuPheAlaValSer 211
 QY 439 AACGTCACTATGCAACAGAGAGAGAGAAATATGACATGAGAGAGTACGCGGAG 498
 DB 212 AlaIleMetTyrLysLleAsnProArgGlnAsnLleProLeuAspSerTher 229
 QY 499 CAGTTTGAAGACTACTCTTCAAGATTTCCTACATCCCGATGAGAGGTGACTCCCT 558
 DB 230 -----GlnAspValProAlaAlaSerValLleValPro 240
 QY 559 TCGCCCTCAGCGAGCC-----CAGCGTGGGAGACACTCTCCCGGAG----- 603
 DB 241 SerAspLeuSerAsnSerValTyrProGlnThrGlyPheTyrSerAsnGlnAspHisLle 260
 QY 604 -----CAA 606
 DB 261 LeuGlnGlyAlaGlyValProSerTyrPheAsnAlaLeuSerValSerAspPheGln 280
 QY 607 GGGCAGCCCT----- 618
 DB 281 GlyTyrAlaGlnThrAlaAlaAsnProValProValPheAlaSerSerLeuProValThr 300
 QY 619 -----GGGGGCACTTCTGAGCCAGACAGATTCATCCCGCTGGGATCCCGGAT 669
 DB 301 HisGlyPheGlyLysSerSerArgSerGlnLleValPhe-----LysValLleCys 318
 QY 670 CCAACCCAGTTTGTGGTGCATCGAAGAAAGAGGCTTGAACATAAGATCATCACT 729
 DB 319 ProLeuCyAsnLleMetArgValLleGlyLysGlyLysThrLleYsArgLleArg 338
 QY 730 AACGAGACCCAGTCCCGGATGATATCCATAGAAAGAGACTCTGAGCTGACAGAG 789
 DB 339 GlnAlaSerLysSerCysLleGlnValAsnAspSerArgThrLysCysGlyAspAspGln 358
 QY 790 CCGTGACATCATGACGCCAGCCAGAGGAGACTTTCAGACGACCGCATGATTCCTGAA 849
 DB 359 CysValLleLleValThrAlaThrGlnSerProAspAspMetLysSerMetAlaValGln 378
 QY 850 -----ATCATGAGAAAGAGGAGATGAGACCAACTAGCCGAGAGATTCTCTGAAA 903

DB 379 AlaValLeuLeuGlnGlyTyrLleAsnAspGlnAspAlaGlnAsnValLysMetGln 398
 QY 904 ATCTGGCACACAAATGCTGTGGAAGACTGTTGAAAAGAGAGAGAGAAATTTGAG 963
 DB 399 LeuValSerSerLysValLleGlyCysValLleGlyLysSerGlySerValLleAsn 418
 QY 964 AAATTCAGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
 DB 419 GlnLleArgLysArgThrAsnAlaAsnLleCysLleSerLysGlyLysAspAspLeu 438
 QY 1018 AGCATATACAAACCCGAGAAAGACCATCTACTGTGAAGGACAGAGTGAAGGCTGCGCAGT 1077
 DB 439 -----ValGlnValSerGlyGlnValSerSerValArgAsp 450
 QY 1078 GCTGAGATGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1125
 DB 451 AlaLeuLleGlnLleValLeuArgLeuArgGlnAspValLleGlyAspLysAspSerVal 470
 QY 1126 -----ATGCTGGCTGTAAACACCCAC 1146
 DB 471 AlaThrArgLysProProAlaArgThrAspAsnCySerPheLeuSerGlySerSerAsn 490
 QY 1147 TCCGGAATC-----TTCTCCAGCTGTATACCCCATCAGCAG 1182
 DB 491 AlaGlyTyrThrLeuProSerPheMetSerSerMetAlaSerThrSerGlyPheHisGly 510
 QY 1183 TTTGGCCCGTCCCG-----CATCATCATCTTAT--- 1212
 DB 511 TyrGlySerPheProAlaGlyAspAsnValLeuGlySerThrGlyProTyrSerThrGly 530
 QY 1213 -----CCAGAGAGAGAGATTTGAACTCTTCAATCCCAACCCAGAGCTGTGGCCCATC 1266
 DB 531 ArgLeuProSerSerSerAlaLeuGlnLleLeuLleProAlaHisAlaMetSerLysVal 550
 QY 1267 ATCGGAG 1326
 DB 551 MetGlyLysGlyGlyGlyAsnLeuGlnAsnLleArgArgLleSerGlyAlaMetLleGln 570
 QY 1327 ATGGCCCTGCGGAG 1386
 DB 571 IleSer---AlaSerLysThrSerHisGlyAspHisLleAlaLeuLeuSerGlyThrLeu 589
 QY 1387 GAA 1389
 DB 590 Gln 590
 RESULT 8
 PCB3 MOUSE STANDARD; PRT; 339 AA.
 ID AC P57722;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Poly(Ic)-binding protein 3 (Alpha-CP3).
 GN PCBp3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP MEDLINE=20396135; PubMed=10936052;
 RX Makeyev A.V., Liehaber S.A.;
 RA Identification of two novel mammalian genes establishes a subfamily
 RT of Xa-domain RNA-binding proteins.";
 RL Genomics 67:301-316(2000).
 CC - FUNCTION: Single-stranded nucleic acid binding protein that binds
 CC preferentially to oligo dC (by similarity).
 CC - SUBCELLULAR LOCATION: Nuclear (by similarity).
 CC - TISSUE SPECIFICITY: Ubiquitous.
 CC - SIMILARITY: Contains 3 KH domains.

Accession	Protein	Length (aa)	MW (kDa)	PI (pI)	Source
KW	Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding				
KW	Repeat.				
FT	DOMAIN	13		KH 1.	
FT	DOMAIN	97		KH 2.	
FT	DOMAIN	261		KH 3.	
SQ	SEQUENCE	339 AA;	35938 MW;	F9F19PF7590CA18	CRC64;

Alignment Scores:	
Pred. No.:	1.49e-07
Score:	234.00
Percent Similarity:	42.12%
Best local Similarity:	25.15%
Query Match:	3.98%
DB:	1
	Gaps:
	12
	339
	Matches:
	83
	Conservative:
	56
	Mismatches:
	11
	Indels:
	74
	12

US-09-270-437D-8 (1-3283) x PCB3_HUMAN (1-339)

QY	643	ATTGATTTCCCGCTGGCGGATTCCTGGTCCGCCACCCAGTTGGTGGTCATCTGGGAAA	702
		::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::	
Db	12	ValThrLeuThrIleLeuArgLeuLeuMetHisGlyLysGluValGlySerIleIleGlyLys	31
		::: :::::::::: :::: :::::::::: :::::::::: :::::::::: ::::::::::	
QY	703	GAGGCGTTGACCATAAAGAACATCATTAAGCAGACCCAGTCCCGGGTAGATATCCATAGA	762
		::: :::::::::: :::: :::::::::: :::::::::: :::::::::: ::::::::::	
Db	32	LysGlyGluThrValLysLysMetArgGluGluSerGlyAlaArgIleAsnIle-----	49
QY	763	AAAGAACTCTGGAGCTGCAGAGAAAGCCTGTCCATCCATGGCCACCACCGAGGGAGCT	822
		50 ---SerGluGlyAsnCysProGluArgIleValThrIleThrGlyProThrAspAlaIle	68
QY	823	TCTGAGAGTCCCGCATGATCTT-----GAAATCATGAGAAAGAGCA	867
		69 PheLysAlaPheAlaMetIleAlaTyrLysPheGluGluAspIleIleAsnMetSer	88
Db	69	PheLysAlaPheAlaMetIleAlaTyrLysPheGluGluAspIleIleAsnMetSer	88
QY	868	GATGAGACCAACTAGCCGAA--GAGATTCCTGAAAATCTTGGACACATAGGCTTG	924
		89 AsnSerProAlaThrSerLysProProValThrLeuArgLeuValAlaProAlaSerGln	108
Db	89	AsnSerProAlaThrSerLysProProValThrLeuArgLeuValAlaProAlaSerGln	108
QY	925	GTTGGAAACATGATTTGAAAGAAAGGCGAATTTGAAAGAAATTTGAACATGAACAGGG	984
		109 CysGlySerLeuIleGlyLysGlyValSerLysIleLysGluIleArgGluSerThrGly	128
Db	109	CysGlySerLeuIleGlyLysGlyValSerLysIleLysGluIleArgGluSerThrGly	128
QY	985	ACCAAGATTAACATTCATCTTTGCGAGATTTGAGCATATTAACAACCCGAAAGACATC	1044
		129 AlaGlnValGlnValAlaGly-----AspMetLeuProAsnSerThrGluArgAlaVal	146
Db	129	AlaGlnValGlnValAlaGly-----AspMetLeuProAsnSerThrGluArgAlaVal	146
QY	1045	ACTGTGAAGGCGACGTTGAGCGCTGTGCCAGTGTGAGATAGAGATATTAAGAAGCTG	1104
		147 ThrIleSerGlyThrProAspAlaIleIleGlnCysValLysGlnIleCysValValMet	166
Db	147	ThrIleSerGlyThrProAspAlaIleIleGlnCysValLysGlnIleCysValValMet	166
QY	1105	CGTGAGGCGC-----	1113
Db	167	LeuGluSerProProLysGlyAlaThrIleProTyrArgProLysProAlaSerThrPro	186
QY	1114	-----TTTGAATAATGATATGCTGGCTGTTAACACCCACTCCGATCTTCCAGCTG	1167
		187 ValIlePheAlaGlyGlyGln-----AlaTyrThrIleGlnGlyGlnTyrAlaIlePro	204
Db	187	ValIlePheAlaGlyGlyGln-----AlaTyrThrIleGlnGlyGlnTyrAlaIlePro	204
QY	1168	TACCCCCCAT-----CACCAATTGGC-----CCGTTCCCGCAT	1200
		205 HisProAspGlnLeuThrLysLeuHisGlnLeuAlaMetGlnGlnThrProPheProPro	224
Db	205	HisProAspGlnLeuThrLysLeuHisGlnLeuAlaMetGlnGlnThrProPheProPro	224
QY	1201	-----CATCACTCTTATCCA	1215
Db	225	LeuGlyGlnThrAsnProAlaPheProGlyGluLysLeuProLeuHisSerSerGluGlu	244
QY	1216	GAGCAGAGATTGTG-----	1233
Db	245	AlaGlnAsnLeuMetGlyLysSerSerGlyLeuAspAlaSerProProAlaSerThrHis	264
QY	1231	AATCTCTTCAATCCCAACCGAGCTGGGGCGCATCATCGGAGAGAGGGGACACATC	1299
		265 GluLeuThrIleProAsnAspLeuIleGlyCysIleIleGlyArgGlnGlyThrLysIle	284

OY 1291 AAACAGGTGGCGATTGCTCCCGAGCGCTCTACATCAAGATTGCCCGCG---GAAAGCCCA 1347
 Db 285 AAGcgtttagtgmtertergyltaglnltlelysllealAaAlarhrgltuglyser 304
 OY 1348 GACGTGAGCGAAGATGCGTATCATCCCGGCGCACCGGAAGCCAGTTCACAGCCNAG 1407
 Db 305 -----SerlgvArglnlmetterlthrltlnrgltYthrProAlaasnllserleuAlaagln 322
 OY 1408 GGAAGGATGCTTGGGAAGACTGAAGAAGGAA 1437
 Db 323 TyrleuileAsnAlargleuThrSerlglu 332
 RESULT 10
 PCB4_HUMAN
 ID PCB4_HUMAN STANDARD; PRT; 403 AA.
 AC P57723;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Poly(rC)-binding protein 4 (Alpha-CP4).
 GN CPB4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=20396135; PubMed=10936052;
 RA Makeyev A.V., Liehaber S.A.;
 RT "Identification of two novel mammalian genes establishes a subfamily
 of KH-domain RNA-binding proteins.";
 RL Genomics 67:301-316(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Isegal T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nagisawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
 RA Takahashi M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
 RA Matanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nishihari K., Masuno Y.,
 RA Niimura K., Iwayanagi T.;
 RA "NEO human cDNA sequencing project.";
 RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Jordan B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marnett K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshimiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.U., Huylk S.W.,
 RA Villalón D.K., Mizny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Single-stranded nucleic acid binding protein that binds
 CC preferentially to oligo dC (By similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -1- SIMILARITY: Contains 3 KH domains.
 CC This SWISS-PROT entry is copyright It is produced through a collaboration

RA Dubois E., el Bakkoury M., Glansdorff N., Messenguy F., Pierard A.,
 RA Schreurs B., Vierendeels F.,
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP PARTIAL CHARACTERIZATION.
 RX MEDLINE=99038243; PubMed=9819425;
 RA Mangus D.A., Amrani N., Jacobson A.;
 RT "Pdp1p, a factor interacting with Saccharomyces cerevisiae poly(A)-
 binding protein, regulates polyadenylation.";
 RL Mol. Cell. Biol. 18:7383-7396(1998).
 CC -1 SUBUNIT: Interacts with PAB1.
 CC -1 SUBCELLULAR LOCATION: Nuclear (Potential).
 CC -1 SIMILARITY: Contains 3 KH domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z36101; CAA85196.1; -.
 CC PIR: S46109; S46109. -.
 CC Germonline: 138776; -.
 DR SGD: S0000437; PBP2.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR Pfam: PF00013; KH; 3.
 DR SMART: SM00322; KH; 3.
 DR PROSITE: PS50084; KH TYPE 1; 3.
 DR KM Nuclear protein; RNA-binding; Ribonucleoprotein; Repeat.
 FT DOMAIN 66 130 KH 1.
 FT DOMAIN 148 213 KH 2.
 FT DOMAIN 330 394 KH 3.
 SQ SEQUENCE 413 AA; 45782 MW; 638846509BCE1840 CRC64;
 Alignment Scores:
 Pred. No.: 3.06e-07 Length: 413
 Score: 229.50 Matches: 105
 Percent Similarity: 42.20% Conservative: 79
 Best Local Similarity: 24.08% Mismatches: 179
 Query Match: 3.90% Indels: 73
 DB: 1 Gaps: 19
 US-09-270-437d-8 (1-3283) x PBP2_YEAST (1-413)
 QY 523 ATTTCTACATCCCGATGAAGAGTGAAGTCCCTTCGCCCCCTCAGAGCCGCGT 582
 Db 10 IletHrThrThrProHrThrValLeuValSerProAsnThrLeuLysArg--LysLys 28
 QY 583 GGGGACCACTCTTCCCGGAGCA----- 606
 Db 29 GtlyGluAspThrSerGluGluGlnLeuGluAlaGluIleLysArgValAlaLeuLysAsp 48
 QY 607 -----GGCCAGGCCCTCGGGGCACTTCTCAGGCCAGACAGATT-----GATTTCCCG 654
 Db 49 AlaAspSerHisSerAspAsnAspHisAspSerProAspAsnValProSerAspValHis 68
 QY 655 CTGCGATCCTGATCCCAACCCAGTTGTGTGTCATCATCGAAAAGAGGGCTTGACC 714
 Db 69 LeuArgMetLeuLysLeuValLysHisAlaSerLeuIleValGlyHisLysGlyAlaThr 88
 QY 715 ATTAAGAACATCACTAAGACAGACCCGCGGTAGATTCATAGAAAAGAACTCT 774
 Db 89 IleserArgIleLysSerGlnThrIleAlaArgIleAsnIleSerAsnAlaLeuArg-- 107
 QY 775 GAGAGCTGAGAGAGAGCTGTCAACATCCATCCAGCCAGAGGAGACTTCTGAAGCATGC 834
 Db 108 GlyValProGlnArgIleValIleValArgGlyThrCysAspAspValAlaLysAlaTyr 127
 QY 835 CGCATGATCTTGAATATCATAGAGAGAGAGAGAGATGAGCAAACTAGCCGAA----- 888

Db 128 GlyMetIleValArgAlaLeuLeuGluGlnHisGlyLysGlnAspAsnGlyLysAla 147
 QY 889 GAGATTCCTCGAAATCTTGACACACATGCTGTTGGAGAAGATTTGAAAAGAA 948
 Db 148 GlnIleSerIleAsnLeuIleIleProHisIleLeuMetGlyCysIleIleGlyLysArg 167
 QY 949 GGCAGAAATTTGAGAAATTTGAACATGAACAGGAGCCAGAACATATCATCTCTTGG 1008
 Db 168 GlySerArgLeuArgGlnIleGluAspLeuSerAlaAlaLysLeuPheAlaSerProAsn 187
 QY 1009 CAGATTTGACATATACACCCGAGAAACCATCATCGAAGAGGAGACAGATTGAGACC 1068
 Db 188 Gln-----LeuLeuLeuSerAsnAspArgIleLeuThrIleAsnGlyValProAspAla 205
 QY 1069 TGTGCGAGTGTGAGATAGATATGAGAAGAGCTGCGTGAAGCCCTTGAATAATGATG 1128
 Db 206 IletHrIleAlaThrPheTyrIleSerGlnThrLeuAsn--PheGlnMetGluSer 224
 QY 1129 CTGGCTGTTAACACC--CACTCCGATATCTC-----TCCAGCCTG 1167
 Db 225 ProGlnLysAsnValLysArgSerIleTyrTyrGlnProThrGlnPheAsnSerValLeu 244
 QY 1168 TACCCCATCACAGTTTGGCCCGCTCCGCAT-----CATCACTTATCCAGAG 1218
 Db 245 IleAspHisSerGlnProAsnThrIlePheHisGlnArgAsnHisGlnIleThrHisProSer 264
 QY 1219 CAGGAGATTGTG-----AATCTCTTATCCCAACCCAGGCTGTGGCCGATC 1266
 Db 265 AspyLysLeuLeuSerTyrLysProAsnLysAsnLeuProIleSerSerThr----- 281
 QY 1267 ATCGGAGAAAGAGGGGACACATCAACAAAGCTGGGG-----AGATTCCCGGAGCTCT 1320
 Db 282 -----LeuLeuSerMetAlaThrProGlnThrThrThrAlaSer 294
 QY 1321 ATCAAGATTGCCCCCGGAGAGGCCAGAC--GTCAAGGAAAGATGATGATCATCATCACC 1377
 Db 295 ValAlaAsnAlaThrAlaPheGlnProAsnPheValIleProAsnValThrValLeuAsp 314
 QY 1378 GGGCCACCGGAGAGCCAGTTCAAGGCCAGAGGAGCGATCTTTGGAAACTGAAAGAGAA 1437
 Db 315 GlyProValIleIleSerProAlaProGlnLysAsnHisLeuMetAsnPheValGlnGlnGlu 334
 QY 1438 AACTCTTTAACCCCAAGAGAAATGAAGTGAAGTGAAGGCGCATATCGAGTCCCTCTCC 1497
 Db 335 IlePheIleAspGluLys----- 340
 QY 1498 ACAGCTCGCGGGGTGTTGGCAAGGTGGCAAGCCGTGAACGAATCGAAGATTAAAC 1557
 Db 341 PheValGlyAsnValIleGlyLysAspGlyLysHisIleAsnSerValLysGluSerThr 360
 QY 1558 AGTGAGAGATCATGCTGCTGCTGTGACCAACGCCAGATGA--AATGAGAAATGATC 1614
 Db 361 GlyCysSerIleIleIle-----GlnAspProValGlnIleLysSerGlnAspArg 377
 QY 1615 GTCAAGATTATGGGCACTCTTTGTTGATCCGACACTGCA-----CAGCGCAAG 1662
 Db 378 LeuThrIleArgGlyThrPheMetAlaSerGlnAlaIleIleMetLeuIleSerAsnLys 397
 QY 1663 ATCAGGGAATGTACACAGGTGAAGCAGAGAGAGAGAGCAAAATATACCT 1710
 Db 398 Ile--GlnIleAspArgSerAsnAlaGluArgLysArgArgSerPro 412
 RESULT 12
 PCBL_HUMAN STANDARD; PRT; 356 AA.
 AC Q15355; Q15157; Q14975;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE POLY(IC)-binding protein 1 (Alpha-CP1) (hmrnp-E1) (nucleic acid
 binding protein SUB2.3).
 GN PCBP1.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9531278; PubMed=7607214;
 RA Leffers H., Dejgaard K., Celis J.E.;
 RT "Characterisation of two major cellular poly(ribo)-binding human
 RL proteins, each containing three K-homologous (KH) domains."; Eur. J. Biochem. 230:447-453 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96016208; PubMed=7556077;
 RA Kiledjian M., Wang X., Liehaber S.A.;
 RT "Identification of two KH domain proteins in the alpha-globin mRNA
 RL stability complex."; EMBO J. 14:4357-4364 (1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lymphocytes;
 RX MEDLINE=94203810; PubMed=8152927;
 RA Aasheim H.-C., Loukianova T., Deggerdal A., Smeland E.B.;
 RT "Tissue specific expression and cDNA structure of a human transcript
 RL encoding a nucleic acid binding [oligo(dC)] protein related to the
 RT pre-mRNA binding protein K."; Nucleic Acids Res. 22:959-964 (1994).
 CC -1- FUNCTION: Single-stranded nucleic acid binding protein that binds
 CC preferentially to oligo dC.
 CC -1- SUBCELLULAR LOCATION: Loosely bound in the nucleus. May shuttle
 CC between the nucleus and the cytoplasm.
 CC -1- TISSUE SPECIFICITY: Abundantly expressed in skeletal muscle,
 CC thymus and peripheral blood leucocytes while a lower expression is
 CC observed in prostate, spleen, testis, ovary, small intestine,
 CC heart, liver, adrenal and thyroid glands.
 CC -1- PM: PHOSPHORYLATED. THE NON-PHOSPHORYLATED FORM(S) EXHIBITED THE
 CC STRONGEST POLY(RC)-BINDING ACTIVITY.
 CC -1- SIMILARITY: Contains 3 KH domains.
 CC -1- CAUTION: There is probably a frameshift error in the nucleotide
 CC sequence of Ref.3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X78137; CAA55016.1; -;
 DR EMBL: U24223; AAA91317.1; -;
 DR EMBL: Z29505; CAA82631.1; ALT_FRAME.
 DR HSSP: Q07244; 1KHM.
 DR Genew; HGNC:8647; PCBP1.
 DR GK; Q15365; -;
 DR MIM; 601209; -;
 DR GO; GO:0005737; C:cytoplasm; NAS.
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003723; F:RNA binding; IDA.
 DR GO; GO:0003697; F:single-stranded DNA binding; IDA.
 DR GO; GO:0016071; P:mRNA metabolism; NAS.
 DR InterPro; IPR004087; KH dom.
 DR InterPro; IPR004088; KH_type_1.
 DR Pfam; PF00013; KH_3.
 DR SMART; SM00322; KH_3.
 DR PROSITE; PS50084; KH_type_1; 3.
 KW Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding;
 KW Phosphorylation; Repeat.
 FT DOMAIN 13 75 KH 1.
 FT DOMAIN 97 162 KH 2.
 FT DOMAIN 279 343 KH 3.
 FT CONFLICT 205 205 V -> A (IN REF. 2 AND 3).
 SQ SEQUENCE 356 AA; 37526 MW; DC85477576DC5104 CRC64;

Alignment Scores:
 Pred. No.: 5.74e-07 Length: 356
 Score: 225.00 Matches: 82
 Percent Similarity: 40.40% Conservative: 61
 Best Local Similarity: 23.16% Mismatches: 123
 Query Match: 3.82% Indels: 88
 DB: 1 Gaps: 11
 US-09-270-437D-8 (1-3283) x PCBP1_HUMAN (1-356)
 QY 619 GGGGCACTTCTCAGGCGCAGACGATTCCTCCGCTGGGAGTCCGTGCCACCCAG 678
 Db 4 GlyValThrGluSerGlyLeuAsnValThrLeuThrIleArgLeuMetHisGlyLys
 QY 679 TTTGTGTGTCATCATCGGAAGAGGCTTGACCAATAAGACATCATCAAGCAGACC 738
 Db 24 GluValGlySerIleIleGlyLysLysGlyGluSerValIleArgGluGluSer 43
 QY 739 CAGTCCCGGTAGATTCATAGAAAAGAACTCTGGAGCTGCAGAGACCTGTCAAC 798
 Db 44 GlyAlaArgIleAsnIle-----SerGluGlyAsnCysProGluArgIleIleThr 60
 QY 799 ATTCATGCGACCCAGAGGGGAGCTTCGAAAGCATGCCGACATTCCTTGAATCATGAG 858
 Db 61 LeuThrGlyProThrAsnAlaIlePheLysAlaPheAlaMetIleLeuAspLysLeuGlu 80
 QY 859 AAGAGGCGAGTAGAGCCAACTA-----GCCGAAGAGATTCTCTG 900
 Db 81 GluAspIleAsnSerMetThrAsnSerThrAlaIleAsnArgProProValThrLeu 100
 QY 901 AAAATCTTGACCAACAATGCTTGTGGAAAGCTGATTTGAAAAGAGCAGAAATTGG 960
 Db 101 ArgLeuValValProAlaThrGlnCysGlySerLeuIleGlyLysGlyCysLysIle 120
 QY 961 AAGAAATTGAACATGAAGAGGAGCAAGTAAACAACTCATCTTGCAGAGATTGAGC 1020
 Db 121 LysGluIleArgGluSerThrGlyAlaGlnValGlnValAlaGly-----AspMetLeu 138
 QY 1021 ATATACACCCGGAAGAACCATCATCTGTGAAGGACAGTTGAGCCTGTGCAGTGTCT 1080
 Db 139 ProAsnSerThrGluArgAlaIleThrIleAlaGlyAlaProGlnSerValThrGluCys 158
 QY 1081 GAGATACAGATT-----ATGAAGAAGCTGCGTAGAGCTTTGAAAATGATATG 1128
 Db 159 ValLysGlnIleCysLeuValMetLeuGluThrLeuSerGlnSerProGlnGlyArgVal 178
 QY 1129 CTGGCTGTAAACCCACCTCCGAGTACTTCTCCAGCTG----- 1167
 Db 179 MetThrIleProGlyGlnProMetProAlaSerSerProValIleCysAlaGlyGln 198
 QY 1168 -----TACCCCAT----- 1176
 Db 199 AspArgCysSerAspAlaValGlyTyrProIleAlaThrHisAspLeuGluGlyProPro 218
 QY 1177 -----CACCAAGTTGGCCCG----- 1191
 Db 219 LeuAspAlaTyrSerIleGlnGlyGlnHisThrIleSerProLeuAspLeuAlaLysLeu 238
 QY 1192 -----TTCCGCCATCATGAC----- 1206
 Db 239 AsnGlnValAlaArgGlnGlnSerHisAspAlaMetMetHisGlyGlyThrGlyPheAla 258
 QY 1207 -----TCTTATCCGAG----- 1218
 Db 259 GlyIleAspSerSerSerProGluValLysGlyTyrThrAlaSerLeuAspAlaSerThr 278
 QY 1213 CAGGAATTTGATGATCTTTCATCCCAAGCCTGTGGCCCATCATCGGAGAGAG 1278
 Db 279 GlnThrThrHisGluLeuThrIleProAsnAsnLeuIleGlyCysIleIleGlyArgGln 298
 QY 1279 GGGGACACATCAAAAGCTGGGAGATTGCGCGAGGCTCTTCAAGANTTGC-----CCT 1335
 Db 299 GlyAlaAsnIleAsnGluIleArgGlnMetSerGlyAlaGlnIleLysIleAlaAsnPro 318

```

QY 1336 GCGGAGGCCGACAGCTCAGCGGAGAAAGATGTCATCATCACCAGGCCACCGGAGCCGAG 1395
Db 319 ValGluGlySer-----SerGlyArgGlnValThrIleThrGlySerAlaIleSerIle 336
QY 1396 TTCAGAGCCGAGGAGCGATCTTGGAGAACTGAAGAGAGAA 1437
Db 337 SerIleuAlaGlnIlyrLeuIleAlaSerIleuSerSerGlu 350

RESULT 13
PCBI_RABIT
ID PCBI_RABIT STANDARD; PRT; 356 AA.
AC 019048;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Poly(rC)-binding protein 1 (Alpha-CPI) (hRNP-B1).
GN PCBP1.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99207122; PubMed=10101190;
RA Thiele B.O., Berger M., Hutch A., Reimann I., Schwarz K., Thiele H.;
RT "Tissue-specific translational regulation of alternative rabbit
RT 15-lipoxygenase mRNAs differing in their 3'-untranslated regions.";
RL Nucleic Acids Res. 27:1828-1836 (1999).
CC -1- FUNCTION: Single-stranded nucleic acid binding protein that binds
CC preferentially to oligo dc (By similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Contains 3 KH domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AJ003023; CA05814.1; -.
DR HSPF: Q07244; 1KHV.
DR InterPro: IPR004087; KH_dom.
DR InterPro: IPR004088; KH_type_1.
DR Pfam: PF00013; KH; 3.
DR SMART: SM00322; KH; 3.
DR PROSITE: PS00084; KH_type_1; 3.
DR Nuclear protein; RNA-binding; Ribonucleoprotein; DNA-binding; Repeat.
KW DOMAIN
FT DOMAIN 13
FT DOMAIN 97
FT DOMAIN 279
FT DOMAIN 343
SQ SEQUENCE 356 AA; 37497 MW; 6D1A261276CA206D CRC64;

Alignment Scores:
Pred. No.: 5.74e-07 Length: 356
Score: 225.00 Matches: 82
Percent Similarity: 40.40% Conservative: 61
Best Local Similarity: 23.16% Mismatches: 123
Query Match: 3.82% Indels: 88
DB: 1 Gaps: 11

US-09-270-437D-8 (1-3283) x PCBI_RABIT (1-356)
QY 619 GGGGGCACTTCTCAGGCGACAGATGATTTCCCGTCGGATCCTGTGTCGCCACCCAG 678
Db 4 GlyValThrGluSerGlyLeuAsnValThrLeuThrIleArgLeuMetHisGlyLys 23
QY 679 TTTGTTGGTCATCATCGGAGAGAGGGCTTGAACATCAATCAATGAAGCAGACC 738
Db 24 GluValGlySerIleIleGlyLysGlyGluSerValIleArgGluGluSer 43

```

```

QY 739 CAGTCCCGGGATGATATCCATGAAAAGAACTCTGAGCTGCAGAGACCTGTACCC 798
Db 44 GlyAlaIleArgLeuAsnIle-----SerGluGlyAsnGlyProGluArgIleIleThr 60
QY 799 ATCCATGCCACCCCGAGAGGAGCTTTCAGACAGCCGCGATGATCTTGAATCATGAGC 858
Db 61 LeuThrGlyProThrAsnAlaIlePheValAlaPheAlaMetIleIleAspIleuGlu 80
QY 859 AAAGAGCGAGTGAACCAACTA-----GCCGAAGAGATTCCTGTG 900
Db 81 GluAspIleAsnSerSerMetThrAsnSerThrAlaIleAspArgProProValThrLeu 100
QY 901 AAAATCTTGCAACAATGCTGCTTGGAAGACTGATGGAAGAAAGAGCAAAATTTG 960
Db 101 ArgLeuValAlaProAlaThrGlnCysGlySerIleuIleGlyLysGlyCysLysIle 120
QY 961 AAGAAATATGAACATGAACAGAGGAGCAAGATTAACAATCTCATCTTTCAGAGATTGAGC 1020
Db 121 LysGluIleArgGluSerThrGlyAlaGlnValGlnIleAlaGly-----AspMetLeu 138
QY 1021 ATATCAACCCGGAAGAACCATCATCTGTGAAGGCGACAGTTGAGGCTGTGCCAGTGT 1080
Db 139 ProAsnSerThrGluArgAlaIleThrIleAlaGlyValProGlnSerValThrGluCys 158
QY 1081 GAGATGAGATTT-----ATGAAGAAGCTGCGGTGAGGCTTTGAAAATGATATG 1128
Db 159 ValLysGlnIleCysLeuValMetLeuGlnThrLeuSerGlnSerProGlnIleArgVal 178
QY 1129 CTGGCTGTAAACCCACCATCGGAGATCTTCACCGCTG 1167
Db 179 MetThrIleProIleGlnProMetProAlaSerProValIleCysAlaGlyGln 198
QY 1168 -----TACCCCAT 1176
Db 199 AspArgCysSerAspAlaAlaGlyTyrProHisAlaThrHisAspLeuGluGlyProPro 218
QY 1177 -----CACCATTTGGCCCG 1191
Db 219 LeuAspAlaTyrSerIleGlnIleGlnIleThrIleSerProLeuAspLeuAlaLysLeu 238
QY 1192 -----TTCCGCAATCATAC 1206
Db 239 AsnGlnValAlaArgGlnGlnSerHisPheAlaMetMetHisGlyGlyThrGlyPheAla 258
QY 1207 -----TCTATCCAGAG 1218
Db 259 GlyIleAspSerSerSerProGlnValLysGlyTyrTrpAlaSerIleuAspAlaSerThr 278
QY 1219 CAGAGATTTGAATCTTTCATCCCAACCCAGCTGTGGGCCCATCTCGGAGAGAG 1278
Db 279 GlnThrThrIleGlnMetThrIleProAsnAsnMetIleCysIleIleGlyArgGln 298
QY 1279 GGGGCGACATCAAAACAGTGGGAGATTCGCGGAGCTCTATCAAGATTCCT 1335
Db 299 GlyAlaAsnIleAsnGlnIleArgGlnMetSerGlyAlaGlnIleLysIleAlaAsnPro 318
QY 1336 GCGGAGGCCGACAGCTCAGCGGAGAAAGATGTCATCATCACCAGGCCACCGGAGCCAG 1395
Db 319 ValGluGlySer-----SerGlyArgGlnValThrIleThrGlySerAlaIleSerIle 336
QY 1396 TTCAGAGCCGAGGAGCGATCTTGGAGAACTGAAGAGAGAA 1437
Db 337 SerIleuAlaGlnIlyrLeuIleAlaSerIleuSerSerGlu 350

RESULT 14
VGLN_HUMAN
ID VGLN_HUMAN STANDARD; PRT; 1268 AA.
AC Q00341; Q9UCY3;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Vglinin (High density lipoprotein-binding protein) (HDL-binding
DE protein).

```



```

QY 961 AAGAAATGTAACATGAACAGGAGCCAGATTAACATTCATCTTTGACGATTGAGC 1020
D 605 LysylserineArginylglutamineAsparagineThreonine-----AspLeuPro 618
QY 1021 ATATACCAACCCG--GAAAGAACCATCTGCTGAAGGCG-----ACAGTTGAGGCC 1068
D 619 AlaGluAsnSerAsnSerGluThrIleIleIleThrGlyLysArgAlaAsnGlyGlu 638
QY 1069 TGTGGCAGTGTGAGATAGAGATTATGAAGAACTGCGCTGAGGCGCTTTGAAATGATG 1128
D 639 AlaArgSerArgIleLeuSerIleGlnLysAspLeuAlaAsnIleAlaGlu----- 655
QY 1129 CTGGCTGTAAACACCCACTCCGGAATCTTCCAGCTGTACCCCATCAGCAGTTGAC 1188
D 655 ----- 655
QY 1189 CCGTTCGCGATCATCTTATTCAGAGCAGAGATTGTGATCTCTTATCCCAACC 1248
D 656 -----ValGluValSerIleProAla 662
QY 1249 CAGCGCTGCGCGCATCATCGGGAAGAGGCGACACATTAACACAGCTGCGCAGATTG 1308
D 663 LysLeuHisAsnSerLeuIleGlyThrLysGlyArgLeuIleArgSerIleMetGluGlu 682
QY 1309 GCCGGAGGCTGTATCAAGATTGCCCTGGCGGAAGGCCAGACGTACGAAAGATGCTC 1368
D 683 CysGlyGlyValHisIleHisPheProValGluGlySer-----GlySerAspThrVal 700
QY 1369 ATCATCAGCGGCGCAGCGGAGCCAGTTCAGAGCCGCGAGCGAGATCTTTGGAAACTG 1428
D 701 ValIleArgGlyProSerSerAspValGluLysAlaLysGluLeuLeu---HisLeu 719
QY 1429 AAGAGGAAACTCTTTTAAACCCCAAGAGAGAGCAAGCTGGAACCGATTCAGAGTG 1488
D 720 AlaGluGlu-----LysGlnThrLysSerPheThrValAspIleArgAla 734
QY 1489 CCCTCTCCACAGCTGCGCGGAGTGGCAAGGTGGCAACCGCTGACGAACTGACG 1548
D 735 LysProGluThrHisLysPheLeuIleGlyLysGlyGlyLysIleArgLysValArg 754
QY 1549 AACTTAACCATGTCAGAACTGCTGCTGTCGACCAAGCCAGATGAAATGAGGAA 1608
D 755 AspSerThrGlyAlaArgValIlePheProAlaAlaGluAspLysAspGln----- 772
QY 1609 GTGATCGTCAAGATTTCGGGCACTTCTTGCACGACACAGTCCAGAGCGAAAGATCAG 1668
D 773 ---LeuIleThrIleLeuGlyLysGluAspAlaValArgGluAlaGlnLysGluLeuGlu 791
QY 1669 GAAATGTGACACAGGTG 1686
D 792 AlaLeuIleGlnAsnLeu 797

```

```

RN [2]
RP FUNCTION.
RX MEDLINE=22120219; PubMed=12124753;
RA Kumar D.V., Nigam A., St John P.A.;
RT "Role of Nova-1 in regulating alpha2N, a novel glycine receptor
RL splice variant, in developing spinal cord neurons.";
RU J. Neurobiol. 52:156-165(2002).
CC -I- FUNCTION: Functions to regulate alternative splicing in neurons by
CC binding pre-mRNA in a sequence-specific manner to activate exon
CC inclusion. It binds specifically to the sequence UCAU (By
CC similarity). Acts to regulate a novel glycine receptor alpha-2
CC chain splice variant (alpha-2N) in developing spinal cord.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- SIMILARITY: Contains 3 KH domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY262017; AAP20872.1; -.
DR InterPro; IPR004087; KH_dom.1.
DR InterPro; IPR004088; KH_type_1.
DR Pfam; PF00013; KH; 3.
DR SMART; SM00322; KH; 3.
DR PROSITE; PS00084; KH_TYPE_1; 3.
KW Antigen; Nuclear protein; RNA-binding; Repeat.
KW BIPARTITE NUCLEAR LOCALIZATION SIGNAL
FT DOMAIN 26
FT FT 48 115
FT DOMAIN 146 212
FT FT 248 384
FT DOMAIN 396 453
FT FT 474 474
FT NON_TER 474
SQ SEQUENCE 474 AA; 48443 MW; 70D0926B5BEC131 CRC64;

Alignment Scores:
Pred. No.: 1,41e-06 Length: 474
Score: 219.50 Matches: 105
Percent Similarity: 35.77% Conservative: 66
Best Local Similarity: 21.97% Mismatches: 190
Query Match: 3.73% Indels: 117
Gaps: 16
DB: 1

US-09-270-437D-8 (1-3283) x NOAL_RAT (1-474)
QY 562 CCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGCAAGCCAGCCCTGGG 621
D 21 ProProAspSerArgLysArgProLeuGlnAlaProProGluAlaGly--SerThrLys 39
QY 622 GGCACCTTCAGGCGCCAGAGATTGATTCCTCCGTCGGATCTGCTCCACCACTTT 681
D 40 ArgThrAsnThrGlyLysGluAspGlyGlnTrpPheLeuValLeuIleProSerThrAla 59
QY 682 GTTGTCGATCATCTGGAAGAGGCGCTTGACCATTAAGACATCACTAAGCAGCCAG 741
D 60 AlaGlySerIleLeuGlyLysGlyGlnThrIleValGlnLeuGlnLysGluThrGly 79
QY 742 TCCCGGTAGATTCATTAAGAAAGAGAC-----TTGGAGCTGCAAGAGCCCTGTC 795
D 80 AlaThrIleLysLeuSerLysSerLysAspPheThrProGlyThrGluArgValCys 99
QY 796 ACATCATCCAGCCAGCCAGGAGGAGCTTCGAGAGCATGCGCATGATCTTGAATCAT 855
D 100 LeuIleGlnGlyThrIleGluAlaLeuAsnAlaValHisGlyPheIleAlaGluLysIle 119
QY 856 CAGAAAGAGCAGATGAGACCAACTAGCCGAAGAGATCTCTG----- 900
D 120 ArgGluMetProGlnAsnValAlaLysThrGluProValSerIleLeuGlnProGlnThr 139

```

Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

Search completed: July 23, 2004, 11:06:57
Job time : 58.8809 secs

```
QY 901 -----AAATCTTGCGACAGAAATGGCTTGTTGGA 930
Db 140 ThrValAsnProAspArgIleYsgInValIysIleIleValProAsnSerThrIaGly 159
QY 931 AGACTGATTGGAAAAGAGCGCAGAAATTGAAAGAAATTGAACATGAACAGGACCAAG 990
Db 160 LeuIleIleGlyLygIlyIaIaThrValIysIaIleIleMerGluInSerGlyIaIaTrp 179
QY 991 ATAAACATCTCATCTTTGACAGATTGACATATACACCCGGAAGAACATCATCTGTG 1050
Db 180 ValGlnIleuSerGlnIlyProAsp--GlyIleAsnIleuGlnGluArgValIaIaThrVal 198
QY 1051 AAGGCGACAGTTGAGCGCTGTGCCAGTGTGAGATAGATATGAGAAAGCGTGGAG 1110
Db 199 SerGlyIuProGluInAsnArgIyIaIaGluIleuIleIleGlnIlyIleGlnIu 218
QY 1111 GCCTTTGAAAATGATATGCTGCTGTAAACACCACTCCGGAATCTTCCAGCCTGTAC 1170
Db 219 -----AspProGlnSerGlySerCysLeuAsnIleSer 229
QY 1171 CCCCATCACAGTTTGCGCGGTCCCGCATCATCAC-----TCTTATCCAGAGCAGAG 1224
Db 230 TyrIaAsnValIThrGlyProValIaIaAsnSerAsnProThrGlyPheProTyrIaAsn 249
QY 1225 ATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGGCGCC--ATCATCGGAGAGAGGG 1281
Db 250 ThrIaGluVal--LeuProThrIaIaIaIaIaIaGlyLeuIleuGlyHis----- 266
QY 1282 GCACACATCAAAACAGCTGGCGAGATTC----- 1308
Db 267 AlaAsnLeuIaGlyValIaIaIaIaPheProIaIaValIeSerGlyPheThrGlyAsnAsp 286
QY 1308 ----- 1308
Db 287 LeuValIaIaIeThrSerIaIeAsnThrIeuaIaSerTyrgIyTyraAsnLeuAsnThr 306
QY 1309 -----GCCGAGCCTCTATCAAGATTGCCCTGGC 1338
Db 307 IeuGlyIeuGlyLeuSerGlnIaIaIaIaIaIaThrGlyIaIeuaIaIaIaIaIaIaSer 326
QY 1339 GAAGGCCCAAGCGTCAGCGAAAGAGTGTATCATCAACCGGCGCACCGAAGCCAGTTC 1398
Db 327 AlaAsnProIaIaIaIaIaIaIaIaAsnLeuIaIaThrTyraIaSerGluIaIa----- 344
QY 1399 AAGGCCAGGAGCGATCTTTGGG----- 1422
Db 345 SerIaSerGlySerThrIaGlyGlyThrIaGlyThrPheIaIeugIySerIeuaIa 364
QY 1423 -----AAACTGAAAGAGAAACTTCTTAACCCCAAGAAAGAGTGAAGCTGGAAGCG 1476
Db 365 AlaIaIaThrIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 384
QY 1477 -----CATATCAGATGCCC 1491
Db 385 IleIeuGlyIThrGluYserThrAspGlySerIyAspValValGluIleIaIaIaValPro 404
QY 1492 TCTTCCACAGCTGGCGCGGTGATTTGGCAAGGTGGCAAGCCGTGAACGAATGCGAGAC 1551
Db 405 GluAsnLeuValIeGlyIaIaIeIeuGlyLygIyIySerThrIeuaIaIuTyrgInGlu 424
QY 1552 TTAAACAGTGCAGAGTATGCTGCTCGT-----GACCAAGCGCAGATGAAGAAATGAG 1605
Db 425 LeuThrGlyIaIaArgIleGlnIleSerIyIyGlyIuPheValIProGlyIThrArgAsn 444
QY 1606 GAAGTATCTGTCAAAATTATCGGCACTTCTTGAACCACTGACAGCTGCACGCGCAAGATC 1665
Db 445 ArgIyIy--ValThrIleThrGlyIThrProIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 463
QY 1666 AGGGAATTTGTACAACAAGTGAAGCAGCGAGCAGAAATACCTCAGGAGATC 1719
Db 464 -----ThrGlnArgIleThrTyrgInGlnIyVal 473
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 23, 2004, 10:55:00 ; Search time 156.961 Seconds

(without alignments)
13198.757 Million cell updates/sec

Title: US-09-270-437D-8

Perfect score: 5886

Sequence: 1 ggcagcgagagagcgagga.....aaccttgaaatgttattc 3283

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=x1d
-O=/cgm2_1/USPFO.spool.P/US09270437/runat_23072004.095739.27425/app.query.fasta.1.5582
-DB=SPRMBL_25 -QFMT=fastan -SUFFIX=fspt -MINMATCH=0.1 -DOOPCL=0 -DOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=DIOSum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=spct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09270437 @CGN 1.1 287 @runat 23072004.095739.27425 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPRMBL_25:*
1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp mhc:*
8: sp organelle:*
9: sp phage:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp unclassified:*
15: sp virus:*
16: sp bacteriophage:*
17: sp archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2838	48.2	556 4 Q9Y6M1	Q9Y6M1 homo sapien

2	2123	36.1	545	11	Q7TOP9
3	1965.5	33.4	594	13	Q73932
4	1965	33.4	593	13	Q57526
5	1951.5	33.2	576	11	Q42254
6	1938	32.9	577	11	Q8CGX0
7	1936	32.9	577	11	Q88477
8	1935	32.9	577	4	Q9NZ18
9	1929	32.8	577	11	Q8BRH1
10	1929	32.8	577	11	Q80US9
11	1890.5	32.1	582	13	Q9PW80
12	1875	31.9	579	11	Q8C2J9
13	1875	31.9	579	11	Q9CPN8
14	1868	31.7	579	4	Q00425
15	863	14.7	580	5	Q81GK4
16	858.5	14.6	573	5	Q81GK4
17	856	14.5	566	5	Q9VZ69
18	478.5	8.1	828	5	Q21605
19	337	5.7	100	4	Q86V81
20	275.5	4.7	641	5	Q9BLA0
21	274	4.7	589	5	Q17935
22	274	4.7	611	5	Q17936
23	272	4.6	568	10	Q9LXF5
24	263.5	4.4	557	5	Q23487
25	261.5	4.4	557	5	Q23487
26	260.5	4.4	169	11	Q7TP50
27	255	4.3	313	4	Q96BP6
28	254.5	4.3	510	5	Q72145
29	254	4.3	680	5	P91393
30	250.5	4.3	653	13	Q7ZXS1
31	250	4.2	398	3	Q74919
32	250	4.2	762	10	Q8LNT9
33	246.5	4.2	833	10	Q9FMF0
34	246	4.2	351	4	Q8N9K6
35	244.5	4.2	1281	13	Q7ZTN2
36	241	4.1	610	13	Q8AX84
37	240	4.1	313	10	Q9LZ82
38	237.5	4.0	313	13	Q7ZYK5
39	237.5	4.0	370	11	Q8BSB0
40	237.5	4.0	542	10	Q857G1
41	237	4.0	371	11	Q85444
42	235	4.0	644	10	Q8LDV1
43	235	4.0	644	10	Q9FNK3
44	234.5	4.0	479	10	Q9X171
45	233.5	4.0	1208	13	Q7ZU44

ALIGNMENTS

RESULT 1
Q9Y6M1 PRELIMINARY; PRT; 556 AA.
ID Q9Y6M1
AC Q9Y6M1;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hepatocellular carcinoma autoantigen.
PE2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99207072; PubMed=10190901;
RA Zhang J.Y., Chan E.K., Peng X.X., Tan E.M.;
RT "A novel cytoplasmic protein with RNA-binding motifs is an autoantigen
in human hepatocellular carcinoma.";
RT J. Exp. Med. 189:1101-1110(1999).
CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF057352; AAD31596.1; -;
DR GO; GO:0005737; Cytoplasm; TAS.
DR GO; GO:0003723; F-RNA binding; TAS.
DR GO; GO:0007345; P-embryogenesis and morphogenesis; TAS.

DR GO: 0009386; P:translational attenuation; TAS.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR InterPro: IPR00504; RNA_rec_mot.
 DR Pfam: PF00013; KH; 4.
 DR Pfam: PF00076; KH; 2.
 DR SMART: SM00322; KH; 4.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS50084; KH_type_1; 4.
 DR PROSITE: PS50102; RRM; 2.
 DR SEQUENCE 556 AA; 61842 MW; 18DEFB10043DCC4 CRC64;

Alignment Scores:

Pred. No.:	4.72e-200	Length:	556
Score:	2838.00	Matches:	556
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	48.22%	Indels:	0
DB:	4	Gaps:	0

US-09-270-437d-8 (1-3283) x Q9Y6M1 (1-556)

QY 70 ATGATGAACAAGCTTTATCATCGGGAACCTGAGCCCGCGCTCACCGCGACGACCTCGG 129
 DB 1 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspAspLeuArg 20
 QY 130 CAGCTCTTTGGGAGACGAAAGCTGCTGCGGGAGACGCTCTGCTGAAGTCCGGCTAC 189
 DB 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40
 QY 190 GCCTTCGAGACTACCCGACGAGAACTGGGCTACCGGCGCATCGACCTCTCGGCT 249
 DB 41 AlaPheValAspTyrProAspGlnAsnTyrAlaIleArgAlaIleGlnThrLeuSerGly 60
 QY 250 AAAGTGAATTCATCGGGAATATCATGGAATTGATTAAGTCTTAAAAAGCTTAAG 309
 DB 61 LysValGlnLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
 QY 310 AGCAGAAATTCAGATTGAAACATCCCTCTCTACCTGACATGGAGAGTGTGGATGGA 369
 DB 81 SerArgLysIleGlnIleArgAsnIleProProHisLysGlnTyrGlnValLeuAspGly 100
 QY 370 CTTTGGGCTCATATCGGACAGTGGAGATGTGGAACAAGTCAACACAGACAGAAACC 429
 DB 101 LeuLeuAlaGlnTyrGlyThrValGlnAsnValGlnGlnValAsnThrAspThrGlnThr 120
 QY 430 GCGGTGTCAAGCTCATATGCAACAGACAGAAAGCAAAATAGCCATGGAGAAAGCTA 489
 DB 121 AlaValValAsnValThrTyrAlaThrArgGlnGluAlaLysIleAlaMetGlnLysLeu 140
 QY 490 AGCGGGCATCAGTTGAGAACTACTCTCTTCAGATTCTCTACATCCCGGATGAAAGGTG 549
 DB 141 SerGlnHisGlnPheGlnLysAsnTyrSerPheLysIleSerTyrIleProAspGlnGluVal 160
 QY 550 AGTCCCTTCGCGCCCTCAGAGGACCCAGCGGTGGGACCACTCTCCGGAGCAAGGC 609
 DB 161 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGlnGlnGly 180
 QY 610 CAGCGCCCTGGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTGCGGATCTGGTC 669
 DB 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
 QY 670 CCCACCCAGTTGTGTGTCATCATCGGAAGAGAGGGCTTGACCTAAAGAACATCACT 729
 DB 201 ProThrGlnPheValGlyAlaIleIleGlyLysGlnGlyLeuThrIleLysAsnIleThr 220
 QY 730 AACGACACCAAGTCCCGGTAGATATCCATGAAAAGAGAACTCTGAGCTGACAGAAAG 789
 DB 221 LysGlnThrGlnSerArgValAspIleHisArgLysGlnLysSerGlyAlaAlaGlnLys 240
 QY 790 CCGTGCACCATCCATGCCACCCGAGAGGGGACTTGTGAAGCATCCGCGCATGATTTTGAA 849
 DB 241 ProValThrIleHisAlaThrProGlnGlyThrSerGlnAlaCysArgMetIleLeuGln 260

QY 850 ATCATGCAAAAGAGGACGATGAGACCAACTAGCCGAGAGATTCCTTGAAATCTTG 909
 DB 261 IleMetGlnLysGlnValAlaAspGlnThrLysLeuAlaGlnLysIleProLeuLysIleLeu 280
 QY 910 GCACACAAATGGCTTGTTGGAAGCTGATTTGGAAAAGAGCGAATTTGAAAGAAAT 969
 DB 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGlnGlyArgAsnLeuLysLysIle 300
 QY 970 GAACATGAACAGGAGCCAAATACATCACTCTTGACAGATTGAGCATATACAC 1029
 DB 301 GlnHisGlnThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
 QY 1030 CCGGAAAGAACCATCACTGTGAGAGGACAGTTGAGGCTGTGCTGAGATAGAG 1089
 DB 321 ProGlnThrGlnThrIleThrValLysGlyThrValGlnAlaCysAlaSerAlaGlnIleGln 340
 QY 1090 ATTATGAAGAGCTGCGTGGAGGCTTTGAAAATGATATGCTGGCTGTAAACACCACTCC 1149
 DB 341 IleMetLysLysLeuArgGlnAlaPheGlnLysAspMetLeuAlaValAsnThrHisSer 360
 QY 1150 GGATTAATTCCTCAGGCTGTACCCCATCACCAGTTGGCCGTTCCGCGCATCACTCT 1209
 DB 361 GlyTyrPheSerSerLeuTyrProHisHisGlnPheGlyProPheProHisHisSer 380
 QY 1210 TATCCAGACGAGAGATTGTGAATCTTTCATCCCAACCGCTGTGGGCCATCATC 1269
 DB 381 TyrProGlnGlnGlnIleValAsnLeuPheIleProThrGlnAlaValAlaIleIle 400
 QY 1270 GGGAAAGAGGGGACACATCAAAAGCTGGCGAGATTGCGCGGACCTCATCAAGATT 1329
 DB 401 GlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIleLysIle 420
 QY 1330 GCGCCGTGGGAGAGGCCCAAGCTCAGCGAAAGATGTCATCATCACCGGCGCACCGGAA 1389
 DB 421 AlaProAlaGlnGlyProAspValSerGlnArgMetValIleIleThrGlyProProGln 440
 QY 1390 GCCCAGTTCAAGGCCCAGGAGCGATCTTTGGAAAATGAAAGAGGAAATCTTTTAA 1449
 DB 441 AlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGlnGlnAsnPheAsn 460
 QY 1450 CCCAAAGAAAGATGAAGCTGGAAGCGCATATCAAGAGGCCCTTCCACAGCTGGCGG 1509
 DB 461 ProLysGlnGlnValLysLeuGlnAlaHisIleArgValProSerSerThrAlaGlyArg 480
 QY 1510 GTGATTTGGAAAGGTGGGAGACCGTGAACGAATGCAAGCTTAACAGTCCAGAGTTC 1569
 DB 481 ValIleGlyLysGlyGlyLysThrValAsnGlnLeuGlnAsnLeuThrSerAlaGlnVal 500
 QY 1570 ATCGTCCCTGTGACCAACGCGCAGATGAAATGAGAAATGATGCTCAGAATTATCGG 1629
 DB 501 IleValProArgAspGlnThrProAspGlnAsnGlnGlnValIleValArgIleIleGly 520
 QY 1630 CACTTCTTTGCTAGCCAGACTCAGACGCGCAAGATCAGGGAATTTACACAGGTGAAG 1689
 DB 521 HisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGlnIleValGlnGlnValLys 540
 QY 1690 CAGCAGAGACAGAAATACCTCAGGGAGTGGCTTCACAGCGAGCAAG 1737
 DB 541 GlnGlnGlnGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 556

RESULT 2
 OTTOF9 PRELIMINARY; PRT; 545 AA.
 AC OTTOF9;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DS Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6; TISSUE=Brain;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold R.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Mang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stopleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,
RA Rata S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Woxley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley K.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC054552; AAH54552.1; -.
KW Hypothetical protein.
SQ SEQUENCE 545 AA; 59679 MW; 8BDD35FF30DF3091 CRC64;

Alignment Scores:
Pred. No.: 1.92e-147 Length: 545
Score: 2123.00 Matches: 418
Percent Similarity: 87.76% Conservative: 12
Best Local Similarity: 85.31% Mismatches: 10
Query Match: 36.07% Indels: 50
DB: 11 Gaps: 2

US-09-270-437d-8 (1-3283) x Q7TQF9 (1-545)
QY 73 ATGAACAAGCTTATACATCGGGAACCTGAGCCCGCGCTGACCGCGGACGACCTCCGCGAG 132
Db 1 MetAsnIysLeuYrIleGlyAsnLeuSerProAlaValThrAlaAspAspLeuAaGln 20
QY 133 CTCTTTGGGAGACGAGAGCTGCCCTGCGGAGACAGGTCTGCTGAAGTCGGGCTACGCC 192
Db 21 LeuPheGlyAspArgIysLeuProIleuAlaGlyGlnValLeuIleuIysSerGlyTyrAla 40
QY 193 TTGCTGAGACTAATCCCGACCGAAGCTGGGCGATCCGGGCGCATGAGACCTCTCGGGTAA 252
Db 41 PheValAspTyrProAspGlnAsnTyrAlaIleArgAlaIleGlnThrLeuSerGlyLys 60
QY 253 GTGGAATTCGATGGGAAATCATGGAAGTATTACCTGCTCTTAAAGCTAAAGAGC 312
Db 61 ValGlnLeuHisGlyIysIleMetGlnValAspTyrSerValSerIysIysLeuAaGser 80
QY 313 AGGAAATTCAGATTGGAACATCCCTCTCACTCAGTCGAGGAGGTGTGTGATGACTT 372
Db 81 ArgArgIleGlnIleArgAsnIleProProHisIleGlnIleThrIleValIleuAspGlyLeu 100
QY 373 TTGGCTCATATATGGGACAGTGGGAATGTGGAACAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuAlaGlnTyrGlyThrValGlnAsnValGlnGlnValAsnThrAspThrGlnThrAla 120
QY 433 GTTGTCAAGCTCAATATGCAACAGAGAAGAAAGAAATATACCTTGGAGAAAGCTAAGC 492
Db 121 ValValAsnValThrIleMetThrArgGlnAlaIysLeuAlaIleGlnIysLeuSer 140
QY 493 GGGCATCACTTTGAGAACTACTCTTCAAGATTCTTCACTCCCGATGAGAGAGTGAAC 552

Db 141 GlyHisGlnPheGlnAspTyrSerPheIysIleSerTyrIleProAspGlnGlnValSer 160
QY 553 TCCCTTCGCCCCCTCAGCGAGCCGACGTG3G3GACCACTTCCCGGAGCAAGGCCAC 612
Db 161 SerProSerProProHisArgAla-----ArgGlnGlnGlyHis 173
QY 613 GCCCTCGGGGACATTTCTCAGCCAGACAGATTGATTTCCCGCTGGCGATCTGTGCTCC 672
Db 174 GlyProGlySerSerSerGlnAlaArgGlnIleAspPheProLeuAaGlyIleValPro 193
QY 673 ACCGAGTTTGTGTCATCATCGAAGAAAGGAGGCTTGACCATTAAGACATCATCATAG 732
Db 194 ThrGlnPheValGlyAlaIleIleGlyIysGlnIleuThrIleIysAsnIleThrIys 213
QY 733 CAGACCAAGTCCCG3GTAGATATCATTAAGAAAGAAAGAACTGTGAGCTGACAGAGCCT 792
Db 214 GlnThrGlnSerArgValAspIleHisArgIysGlnIysSerGlyAlaIleGlnIysPro 233
QY 793 GTCAACCATTCATGCCACCCCGAGAGGACCTTCTAAGCATGCGCATGATTTCTGAATC 852
Db 234 ValThrIleHisAlaThrProGlnGlyThrSerGlnIaCysArgMetIleLeuGlnIle 253
QY 853 ATGCAGAAAGAGCGAGATGAGACCAAACTAGCGGAGAGATTCCTTGAAATCTTGCA 912
Db 254 MetGlnIysGlnIleAspGlnThrIleIysLeuAlaGlnGlnValProLeuIysIleLeuAla 273
QY 913 CACAATGCTGTGGTGAAGACTGATTTGAAAGAAAGAGAGAAATTTGAAGAAATTTGAA 972
Db 274 HisAsnGlyPheValGlyArgLeuIleGlyIysGlnIysArgIleuIysIleGlnIle 293
QY 973 CATGAACAGAGAACCAAGATTAACAATCTCATCTTTGAGAGATTTGAGCATATACACCG 1032
Db 294 HisGlnThrGlyThrIysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAspPro 313
QY 1033 GAAAGAAACATCACTGTTGAAGGCGACAGTGTAGGCGCTGTGCGATGCTGATGAGATT 1092
Db 314 GlnArgThrIleThrValAlaGlyIleThrIleGlnIaCysAlaAsnAlaGlnIleGlnIle 333
QY 1093 ATGAAGAAGCTGCGTGAAGGCTTTGAAATGATATGATGCTGCTGTAAAC----- 1140
Db 334 MetIleIysLeuArgIleGlnIlePheGlnAsnAspMetLeuAlaValAsnGlnAlaAsn 353
QY 1140 ----- 1140
Db 354 LeuIleProIleuAsnIleuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 373
QY 1141 -----ACC 1143
Db 374 ProProProAlaGlyProArgGlyValProProSerProProTyrHisProPheAlaThr 393
QY 1144 CACTCGGATTACTTTCAGCCCTGTACCCCATCACAGTTTGCCCGCTCCGCGATCAT 1203
Db 394 HisSerGlyTyrPheSerSerLeuTyrProHisIleHisPheGlyProPheProHisHis 413
QY 1204 CACTTTATTCAGAGACGAGATTTGAAATCTTTATCCCAACCAAGGCTGTGGCGCC 1263
Db 414 HisSerTyrProGlnIleGlnThrValSerLeuPheIleProThrGlnAlaValGlyAla 433
QY 1264 ATCATGGGAAAGAGGGGCGACACATCAAAACGCTGCGGAGATTCGCGGAGCTCATG 1323
Db 434 IleIleGlyIysIysGlyAlaHisIleIysGlnIleuAlaArgPheAlaGlyAlaSerIle 453
QY 1324 AAGATTGCCCTTCGCGAAGGCCGACAGTCAAGGAAAGAGTGTCAATCACCGGAGCA 1383
Db 454 LysIleAlaProAlaGlnIleGlyProAspValSerGlnArgMetValIleIleThrGlyPro 473
QY 1384 CCGGAAGCCCAAGTTCAAGGCCCAAGGAGCG 1413
Db 474 ProGlnAlaGlnPheIysValSerAlaIys 483
RESULT 3
073932
ID 073932 PRELIMINARY; PRT; 594 AA.

AC 073932; (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, last annotation update)
 DE vgl RNA binding protein variant D.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 CC Xenopodinae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
 RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
 RL Genes Dev. 0:0-0(1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98228351; PubMed=9560341;
 RA Deshler J.O., Hight M.I., Adameson T., Schnapp B.J.;
 RT "A highly conserved RNA-binding protein for cytoplasmic mRNA
 localization in vertebrates."
 RL Curr. Biol. 8:489-496(1998).
 CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AF064634; AAC18598.1; -;
 DR EMBL; AF055923; AAC41285.1; -;
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH_dom.
 DR InterPro; IPR004088; KH_type_1.
 DR InterPro; IPR005054; RNA_rec_mot.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PSS0084; KH_type_1; 4.
 DR PROSITE; PSS0102; RRM; 2.
 SQ SEQUENCE 594 AA; 65643 MW; 5ACEAYBF0856D6 CRC64;
 Alignment Scores:
 Pred. No.: 7,68e-136 Length: 594
 Score: 1965.50 Matches: 388
 Percent Similarity: 78.50% Conservative: 72
 Best Local Similarity: 66.21% Mismatches: 91
 Query Match: 33.39% Indels: 35
 Db: 13 Gaps: 5
 US-09-270-437D-8 (1-3283) x 073932 (1-594)
 QY 73 ATGAACAAGCTTTTAACTCGGAACTGAGCCCGCGTACCGCGAGCACTCCGCGAG 132
 Db 1 MetAshnlyLeuTyrTlleglYasnleuSerGluAenValSerProThrAspLeuGlnSer 20
 QY 133 CTCTTTGGGACAGAGAGCTGCCCTGGCGGACAGAGCTCTGTGAAGTCCGGCTACGGC 192
 Db 21 LeuPhelyGlnSerYsilleProPhethrGlnPhleuVallysserGlyTyrAla 40
 QY 193 TTGCTGAGTACCCGACGAGAACTGGGCCATCGGCCATCGAGACCTCTCGGCTAA 252
 Db 41 PheValAspCysProAspGlnuTrpAlaMetLysAlaTllespThreusSerGlyLys 60
 QY 253 GTGGAATTGCATGGGAAATCATGAAAGTTGATTACTCACTCTTAAAAAGCTAAGAGC 312
 Db 61 ValGlnLeuHlsGlyLysValTlleglValGlnHlsSerValProLysAspGlnAspSer 80
 QY 313 AGGAATAATTCAGATTGGAATCATCCCTCTCTACCTGACGTGGAGAGGTGTGATGACTT 372
 Db 81 ArgLysLeuGlnTlleglYasnTlProProHlsLeuGlnTlPrpLlValLeuAspSerLeu 100
 QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGGACAAGTCAACAGACAGACAGAACCGCC 432
 Db 101 LeuAlaGlnTyrTlThrValGluAsnCysGlnGlnValAsnThrGlnSerGlnTlPrpAla 120
 QY 433 GTTGCAAGCTCACATATGCAACAGAGAAAGAAATATGCAATGAGGAAAGTAAAGC 492
 Db 433 GTTGCAAGCTCACATATGCAACAGAGAAAGAAATATGCAATGAGGAAAGTAAAGC 492

Db 121 ValValAsnValThrTyrAlaAsnLysGlnHlsAlaArgGlnGlyLeuGlnLysLeuAsn 140
 QY 493 GGGCATGAGTTTGAAGACTACTCTCAAGATTCTTACATCCCGAGTGAAGAGTGAAGC 552
 Db 141 GlyTyrGlnLeuGlnAsnTyrSerLeuLysValThrTyrTlProAspGlnMetAlaThr 160
 QY 553 TCCCTTCGCCCCCTCAGCGAGCCAGCGGGAGCACTTCCCGGAGACAGGCCAC 612
 Db 161 ProGlnAlaProSerGlnGlnLeuGlnGlnProGlnGlnGlnHlsPProGlnGlyArg 180
 QY 613 -----GCCCTGGGGCGACTTCGAGCC 636
 Db 181 ArgLysPheGlnArgLysProAlaArgGlnGlySerProGlyAlaAlaAlaArgPro 200
 QY 637 AGA---CAGATTGATTTCCCGCTGCGAGTCTGTCCTCCACCACTTTGTTGGTGCATC 693
 Db 201 LysProGlnThrGlnValProLeuArgMetLeuValProThrGlnPheValGlyAlaTle 220
 QY 694 ATCGGAAGAGGGCTTGACCATTAAGACATCACTTAAGCAGACCCAGTCCCGGTAGAT 753
 Db 221 TlleglYsGlnGlyAlaThrTlleglYasnTlThrLysGlnThrGlnSerLysTleAsp 240
 QY 754 ATCCATGAATAAGAACTCGGAGCTGACAGAAAGCTGTACACCATCCATGCCACCCA 813
 Db 241 TllesAspGlnGlnAsnAlaGlyAlaAlaGlnLysProTlThrTllesHlsSerThrPro 260
 QY 814 GAGGGACCTTGGAAGCATCGCGCATGATCTTGAAATCATGACAGAAAGGCAGATGAG 873
 Db 261 GlnGlyCysSerAlaAlaCylLysTlTleMetGlnTlMetGlnGlnGlnAlaGlnAsp 280
 QY 874 ACCAAATGAGCCGAAGAGATTCCTTGAATCTTGGCACACAAAGGCTTGTTGGAAGA 933
 Db 281 ThrLysPheThrGlnGlnTlProLeuLysTleuValAsnAsnAsnPheValGlyArg 300
 QY 934 CTGATTGAAAAGAGGAGAAATTTGAAGAAATTTGAACATGAACAGAGCAAGCAAGATA 993
 Db 301 LeuTlleglYsGlnGlyArgAsnLeuLysTlleglGlnAsnPhrThrAspThrLysTle 320
 QY 1054 GGCACAGTTGAGGCTGTGCCAGTGTGATAGAGATTATTAAGAGCTGGGTGAGGCC 1113
 Db 341 GlySerTlleglProCysAlaLysAlaGlnGlnTlMetLysLysTleArgGlnSer 360
 QY 1114 TTTGAAAATGATATGCTGGCTGTTAACACCACTCC----- 1149
 Db 361 TyrGluAsnAspTlTleAlaMetAsnLeuGlnAlaHlsLeuTlProGlyLeuAsnLeu 380
 QY 1150 -----GATACCTTC-----TCAGCTGTGACCCCATGACCACTTTGGC--- 1188
 Db 381 AsnAlaLeuGlyLeuPheProSerSerSerGlyMetProProProSerValGlyVal 400
 QY 1189 -----CCGTCCCGCATCATCACTCTTAATCCAGAGCAGAGATT 1227
 Db 401 ProSerProThrSerSerThrSerTyrProProPheGlyGlnGlnProGlnSerGlnTlPr 420
 QY 1228 GTGAATCTTTCATCCCAACCAAGAGCTGTGGCGGCATATCGGAGAAAGGGGACAC 1287
 Db 421 ValHlsLeuPheTlProAlaLeuAlaValGlyAlaTlTleLysGlnGlnGlnHls 440
 QY 1288 ATCAACAGCTGGCGAGATTGGCGGAGCTCTATGAAGATTGCGCTGGAGAGGCCCA 1347
 Db 441 TlleglGlnLeuSerArgPheAlaGlyAlaSerTlLysTlLysAlaProAlaGlnGlyPro 460
 QY 1348 GACGTACGGAAGAGATGCTCATATCAACCGGGCCACCGGAACCGGATTCAGAGGCCAG 1407
 Db 461 AspAlaLysLeuArgMetValTlTlTlThrGlyProProGlnAlaGlnPheLysAlaGln 480
 QY 1408 GAGCGAGATCTTGGGAAAGTGAAGAGAAAGAACTCTTTAAACCCCAAGAGAGTGAAG 1467
 Db 481 GlyArgTlTyrTlGlyLysLeuLysGlnGlnAsnPhePheGlyProLysGlnGlnValLys 500

QY 1468 CTGGAAGCCGATATGAGATGCGCTCTTCCAGAGCTGGCGGGATGATTGGCAAGTGGC 1527
 Db |||||
 Db 501 LeuGluThrHisIleValProSerTyrIalaIleValIleGlyIleGly 520
 QY 1528 AAGACCGTGAAGCACTGAGACTTAAACAGTGAAGAGTATCGTGGCTGGACCA 1587
 Db |||||
 Db 521 LysThrValAsnGluLeuGlnAsnLeuThrSerIleGluValIleValProArgAspGln 540
 QY 1588 ACGCAGATGAAATAGGAGATGATCGTCAAGATTATCCGGGACTTCTTGTGCTAGCCAG 1647
 Db |||||
 Db 541 ThrProAspGluAsnAspIleValIleValIleValIleGlyHisPheTyrIleSerGln 560
 QY 1648 ACTGCACAGCGCCAGATCAGGAAATTTGTACACAGTGAAGCAGACGAGCAAGAAATAC 1707
 Db |||||
 Db 561 LeuIleGlnArgGlyIleGlnIleLeuIleGlnValIleArgGlnGlnGlnGln 580
 QY 1708 CCTCAGGAGTGGCTCA 1725
 Db |||||
 Db 581 GlnLysThrValGlnSer 586

RESULT 4

ID 057526 PRELIMINARY; PRT; 593 AA.

DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE KH domain-containing transcription factor B3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92249652; PubMed=1577195;
 RX Pfaff S.L., Taylor W.L.;
 RA "Characterization of a Xenopus oocyte factor that binds to a
 RT developmentally regulated cis-element in the TFIID gene";
 RL Dev. Biol. 151:306-316(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Griffin D., Taylor W.L.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
 RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
 RL Genes Dev. 0:0-0(1998).
 CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL: AF042353; AAB97457.1; -
 DR EMBL: AF046433; AAC18597.1; -
 DR GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro: IPR004087; KH_dom.
 DR InterPro: IPR004088; KH_type_1.
 DR InterPro: IPR005054; RNA_rec_mot.
 DR Pfam: PF00013; KH; 4.
 DR Pfam: PF00076; rtm; 2.
 DR SMART: SM00322; KH; 4.
 DR SMART: SM00360; RRM; 2.
 DR PROSITE: PS50084; KH_type_1; 4.
 DR PROSITE: PS50102; RRM; 2.
 DR PROSITE: PS00030; RRM_RNP_1; FALSE NEG.
 SQ SEQUENCE 593 AA; 65385 MW; 5A5AB4B4A1D5DF7 CRC64;

Alignment Scores:

Pred. No.: 8.35e-136 Length: 593
 Score: 1965.00 Matches: 388
 Percent Similarity: 79.27% Conservative: 71
 Best Local Similarity: 67.01% Mismatches: 84
 Query Match: 33.38% Indels: 36
 Db: 13 Gaps: 6

US-09-270-437d-8 (1-3283) x 057526 (1-593)

QY 73 ATGAACAAGCTTTATCATCGGGAACCTTGAGCCCGCGTACCGCCGACGACTCCGGCAG 132
 Db |||||
 Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGlnValSerProProAspLeuGluSer 20
 QY 133 CTCTTTGGGAGCAGAGAGCTCCCTGGCGGAGCAGTCTTGCTGAGTCCGGCTACGCGC 192
 Db |||||
 Db 21 LeuPheLysGluSerLysIleProPheThrGlnPheLeuValLysSerGlyTyrIle 40
 QY 193 TTGCGGACCTACCCCGACCAAGAACTGGGGCCATCCGGCCATTCGAGACCTCTCGGGTAAA 252
 Db |||||
 Db 41 PheValAspCysProAspGluThrTyrIleMetLysAlaIleAspThrLeuSerGlyLys 60
 QY 253 GTGGAATTCATGGGAAATATCATGATTATTAATCACTGCTCTTAAAGTAAGTACAGGC 312
 Db |||||
 Db 61 ValGluLeuHisGlyLysValIleGluValGlnHisSerValProLysArgGlnArgSer 80
 QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCCTCAGTGAAGTGGAGTGTGATGACTT 372
 Db |||||
 Db 81 ArgLysLeuGlnIleLeuArgAsnIleProProHisLeuGlnIleProGluValLeuAspSerLeu 100
 QY 373 TTGCGTCAATATGGGACAGTGAAGATGTGGAATGTCAACAAGACACAGAAACCGCC 432
 Db |||||
 Db 101 LeuIleGlnTyrGlyThrValGluAsnCysGlnIleValIleAsnThrAspSerGluThrAla 120
 QY 433 GTTGTCACGTCATCATATGCAACAAAGAAAGAAAGAAATAGCATGAGACGACTAGC 492
 Db |||||
 Db 121 ValValAsnValThrTyrIleAsnLysGlnHisIleAlaArgGlnGlyLeuGluLysLeuAsn 140
 QY 493 GGGCATCATGTTTGAAACTACTCTTCAGATTTCTTACATCCCGGATGGAAGATGAGC 552
 Db |||||
 Db 141 GlyTyrGlnLeuGlnValSerLeuLysValThrTyrIleProAspGluMetAlaThr 160
 QY 553 TCCCTCTGGCCCTCGACGACCCACCGT----- 582
 Db |||||
 Db 161 ProGlnSerProSerGlnGlnLeuGlnIleProGlnGlnHisProGlnGlyArgArg 180
 QY 583 -----GGGACCACTCTCCCGGAGCAGAGCCAGCCCTGGGGGCACTTCCACAGCC 636
 Db |||||
 Db 181 GlyPheGlyGlnArgLysProAlaArgGlnGly---SerProGlyAlaIleAlaAspPro 199
 QY 637 AGA---CAGATTGATTTCCCGTGGCGAGATCGTCCCAACCAATTTGTTGGTCCATC 693
 Db |||||
 Db 200 LysProGlnSerGlnValIleProLeuArgMetLeuValProThrGlnPheValGlyAlaIle 219
 QY 694 ATCGAAAGGAGGGCTGACCAATTAAGAAATCATTAAGACAGCCAGTCCCGGTATGAT 753
 Db |||||
 Db 220 IleGlyLysGlnGlyAlaThrIleArgAsnIleThrLysGlnThrGlnSerLysIleAsp 239
 QY 754 ATCCATAGAAAGAGAACTCTGAGCTGCAGAGAGCCGTGCATCCATGCGACCCCA 813
 Db |||||
 Db 240 IleHisArgLysIleAsnIleValIleAlaGluLysProIleThrIleHisSerThrPro 259
 QY 814 GAGGGGACTTCTGAGCATGCGCATGATTTCTGAATCATGACGAAAGAGCAGATGAG 873
 Db |||||
 Db 260 GluGlyCysSerAlaIleCysLysIleIleMetGlnIleMetGlnLysGluAlaGlnAsp 279
 QY 874 ACCAACTAGCCGGAAGAAGATTCCTCTGAAATTTGTGGCACACATGGCTTGTGAGACA 933
 Db |||||
 Db 280 ThrLysPheThrGlnIleIleProLeuLysIleLeuAlaHisAsnAspPheValGlyArg 299
 QY 934 CTGATTGGAAGAAAGAGCAGAAATTTGAGAAATGAGATCAATGAACGAGGACCAAGATA 993
 Db |||||
 Db 300 LeuIleGlyLysGlnGlyArgAsnLeuLysIleGlnGlnAspThrAspThrLysIle 319
 QY 994 ACAATTCATCTTTCAGGATTTGACATATACCAACCCGAAAGAAACCATCACTCTGAGAG 1053
 Db |||||
 Db 320 ThrIleSerProLeuGlnIleAspLeuThrLeuTyrAsnProGluArgThrIleThrValLys 339
 QY 1054 GGCACAGTTGAGGCGCTGGCCAGTGTCTGAGATAGAGATTTAGAAAGCTGGGTGAGGCC 1113
 Db |||||

Db	340	GlySerIleGluThrCysAlaIysAlaGluGluValMetIysLysIleArgGluSer	359
QY	1114	TTTGAATAATATATGCTGCTGTTAAACACCACTCC-----	1149
Db	360	TyrGluAsnAspIleAlaAlaMetAsnLeuGlnAlaHisIleuIleProGlyLeuAsnLeu	379
QY	1150	-----GGATACCTTC-----TCCAGCTGTACCCCATCACCAATTGGC---	1188
Db	380	AsnAlaLeuGlyLeuPheProProSerSerSerGlyMetProProProSerAlaGlyVal	399
QY	1189	-----CCGTCGCCGATCATCACTCTTATCCAGACGACGAAATT	1227
Db	400	SerSerProThrThrSerSerAlaSerTyrProProPheGlyGlnGlnProGluSerGluThr	419
QY	1228	GTGAATCTCTTCATCCCAACCCAGGCTGTGGCGCCATCATCGGAAGAAAGGGGACAC	1287
Db	420	ValHisIleuPheIleProAlaLeuAlaValGlyAlaIleIleGlyLysGlnGlyHis	439
QY	1288	ATCAAAACAGCTGGCGGAATTGCGCGAGCCTCTTATCAAGATTGGCCCTCGCGGAAGGCCCA	1347
Db	440	IleGlyGlnLeuSerArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluGlyPro	459
QY	1348	GACGTCAACGAAAGATGGTGATCATATACCGGGGACACCGAAGCCCAATTCAAGGCCAG	1407
Db	460	AspAlaIysLeuArgMetValIleIleThrGlyProProGluAlaGlnPheLysAlaGln	479
QY	1408	GGACGGATCTTTGGGAAACTGAAGAGGAAACTTCTTTAACCCCAAGAAAGACTGAAG	1467
Db	480	GlyArgIleTyrGlyLysLeuLysGluLysAsnPheGlyProLysGlyGluValLys	499
QY	1468	CTGGAAGCGCATATCATGAGTGCCCTCTTCCACAGCTGGCCGGGTGATTGGCAAGCTGGC	1527
Db	500	LeuGluAlaHisIleLysValProSerTyrAlaAlaGlyArgValIleGlyLysGlyLys	519
QY	1528	AAGACCGTGAACGAACTGCAAACTTAACAAGTGAAGAAAGTCACTGCTGTACCAAA	1587
Db	520	LysThrValAsnGlyLeuGlnAsnMetThrSerAlaGluValValProArgLysArgLys	539
QY	1588	ACGCAGATGAAATAGAGAAATGCTGCTCAGAAATTATCGGGCACTTCTTGCTTACCCAG	1647
Db	540	ThrProAspGluAsnAspGlnValValLysIleThrGlyHisPheTyrAlaSerGln	559
QY	1648	ACTGCACAGCCGCAAGATCAGGAAATTGTACAACACGCTGACGACAGCAGACGAGAA	1704
Db	560	LeuAlaGlnArgLysIleGlnGluIleLeuAlaGlnValArgArgGlnGlnGln	578
RESULT 5			
ID	042254	PRELIMINARY;	PRT; 576 AA.
AC	042254;		
DT	01-JAN-1998	(TREMBLrel. 05, Created)	
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Zipcode-binding protein.		
GN	ZBP1.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
RN	[1]-----		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=97220007; PubMed=9121465;		
RA	Ross A.F., Olevnikov Y.S., Kislauksis E.H., Taneja K.L., Singer R.H.,		
RT	"Characterization of a beta-casein mRNA zipcode-binding protein.";		
CC	-1- SIMILARITY: CONTAINS 4 KH DOMAINS.		
DR	EMBL; AF026527; AAB82295.1; "		
DR	GO; GO:0003676; F: nucleic acid binding; IEA.		
DR	InterPro; IPR004087; KH dom.		
DR	InterPro; IPR004088; KH_type_1.		
DR	InterPro; IPR000504; RNA_rec_mot.		
DR	Pfam; PF00013; KH; 4.		

[illegible]


```

QY 898 CTGAAATCTTGCGACACAAATGGCTTGTTGGAGACTGATTGAAAAAGAGCAGAAAT 957
Db 279 Leuylsileuvalahisasmnshphevalglagluileglylsuylarasn 298
QY 958 TTGAAGAAATTTGAACATGAAACAGGAGCAGGATTAACATCTCTTTGGCAGATTGG 1017
Db 299 Leuylsvalaglunaspthrighrlystlethrlyleaserleuglnaspheu 318
QY 1018 AGCAATATCAACCCGGGAAAGAACCTCACTGCTGGAAAGGCGACAGTTCAGGCTTCACAGT 1077
Db 319 Thrleuylasmproglnarghrlythrlyvalysgllyserlleghlasmcysyls 338
QY 1078 GCTGAGATAGAGATTATGAAAGAGCTGCGTGAAGGCTTGAATGATGCTGCTT 1137
Db 339 Alagluingluilemetlylsvalarglulalatyrglulasmavalalalamet 358
QY 1138 AACACCACTCC-----GGATCTTC----- 1158
Db 359 SerleuglnserhstleulleProglYleuasnleulalavalglYleupheProala 378
QY 1159 -----TCCAGCTGTACCCCATACCAG 1182
Db 379 SerserasnlaValProProProProSerservalserglYalalalProlyrSerser 398
QY 1183 TTTGGCCCGTTCGCGCATCATCTTATCCAGAGCAGAGATTGGATCTCTTCATC 1242
Db 399 PhehetPro-----Proglungluinrvalhstvalshelle 411
QY 1243 CCAACCAAGGCTGTGGCGCCATCATCGGAAAGAGGCGCACATCAACAGCTGGCG 1302
Db 412 ProalaglnalavalglYalallelglYlyslysglYnhstlelysglnleuser 431
QY 1303 AGATTGCGCGGAGCTCTATCAAGATTGCCCCCTGCGGAAAGCCGACGTCAGCGAAAG 1362
Db 432 ArpPhehlaserhstlelystlealaproProglYthrProaspsrlylsValarg 451
QY 1363 ATGTCATCATCAACCGGCGCACCGAGCCAGTTCAGGCGCCAGGAGCGATCTTGGG 1422
Db 452 MetValVallethrghYpropoglnalaglnpheylsalaglnlYatgltYrghY 471
QY 1423 AAACTGAAAGAGAAAATTCTTTAACCCCAAGAAAGTGAAGCTGGAAGCGCATATC 1482
Db 472 LysleuylsgluYlunshpheghYProlysglnluValYlsuengluYthrValaengl 491
QY 1483 AGAGTGCCCTCTCCACAGCTGCGCGGATGTTGGCAAGCTGGCAAGACCGTGACGA 1542
Db 492 ArgvalProalaserhstlealaglYargvalilleglYysglYglYsthrValaengl 511
QY 1543 CTGCAAACTTAACAGATGAGAGATGATGCTGCTGTCGCAACGCGCCAGTTGAAAT 1602
Db 512 LeuglnasheuthralaahagluvalValValProargaspghnhrProaspsgluasn 531
QY 1603 GAGGAAGTATGCTCAAGATTATCGGCACTTCTTGTCTAGCAGACTGACAGCCGCAAG 1662
Db 532 GlunghlvalillevalylsleilleglYhstpheYlaserghlnmetahaglnaglyls 551
QY 1663 ATGAGGGAATTTTACAACAGGTGAAGCCAGCAGAGCAGAAATACCTCAGGAGTGGCC 1722
Db 552 IleargAspilleuvalaglnvalYlsghlnghlnhstlelysglnserghYlnleu 571
QY 1723 TCACAGCGCAGCAAG 1737
Db 572 GlnalargArglyls 576

```

```

RESULT 6
Q8CGX0
AC Q8CGX0; PRELIMINARY; PRT; 577 AA.
DT 01-MAR-2003 (Tremblurel. 23, Created)
DT 01-MAR-2003 (Tremblurel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblurel. 23, Last annotation update)
DE B-actin zipcode binding protein 1.

```

```

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN=Sprague-Dawley;
RA Eom T., Singer R.H., Bassell G.J.;
RT "Molecular interactions between rZBP1 and b-actin zipcode required for
RT transport of mRNA and stimulation of spine growth.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF541940; A016210.1;
DR GO: 0003676; P:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63436 MW; 0647676128FBD1BE CRC64;

```

Alignment Scores:

```

Pred. No.: 8.03e-134 Length: 577
Score: 1938.00 Matches: 382
Percent Similarity: 79.10% Conservative: 76
Best Local Similarity: 65.98% Mismatches: 95
Query Match: 32.93% Indels: 26
DB: 11 Gaps: 5

```

US-09-270-437D-8 (1-3283) x Q8CGX0 (1-577)

```

QY 73 ATGACAAGCTTTACATCGGGAACCTGAGCCCGCGTCAACCGGACGACTCCGCGAG 132
Db 1 MetAsnlylsleuylrilleghYasnleuAsnlyserValthrProalaaspleuYlulys 20
QY 133 CTCTTTGGGACAGAGAGCTGCCCCCTGCGCGGACAGAGTCTGCTGAAGTCCGCTACGCC 192
Db 21 ValPhealaglnhstlystleaserlyrserghYghlnpheuvalYlsrserghYrghY 40
QY 193 TTGTGTGACTACCCCGACCAAGACTGGCGCATTCGCGCATGAGACCTTCGCGTAA 252
Db 41 PheValAspCysProaspghlnhstlyrPalametlysalilleglYthrPheaserghYls 60
QY 253 GTGCAATTGATGAGGAATTCATGGAAGTTGATTACTCACTCTTAAAGCTAAGAGAC 312
Db 61 ValghluenglnlelylsargluengluilleghlnhstleValProlylsYsglnargser 80
QY 313 AGGAAATTCAGATTGGAACATATCCCTCTCACTGCGAGTGGAGGTGTGATGACTT 372
Db 81 ArglylilleghlnleargsnlleProProglInleuarghtrghlValleuaspserleu 100
QY 373 TTGGCTCATATTGGGACAGTGAAGATTGGAAACAAGTCAACAGACAGAACCGCC 432
Db 101 LeuValaglnlyrghYthrValghluenglnCysghlnghlnValasnthrghluserghYthrAla 120
QY 433 GTTGTCAAGTCAATATGCAACAGAGAAAGAAAGCAAAATGCCATGAGAAAGCTTAAGC 492
Db 121 ValValasnValThrlyrserasnhrghYlunghYthrArgghlnalalemetlylsleuasn 140
QY 493 GGGCATCAGTTTGAGAACTACTCTTCAAGATTCTTACATCCCGGATGAAGAGTGAGC 552
Db 141 GlnhstleghlnleughlnhstlysalaleuYsvalserlylleProaspsglunilleal 160
QY 553 TCCCTTCGCCCTTCAGCGGACCGACGCTGGGAGACACTTTCCCGGAGAGAGCCAC 612
Db 161 -----GlnghYProglunasnghYargargghYlghYrghYserghYlghYlnProarg 178
QY 613 -----GCCCTGGGGGCACTTTCAGGCGCAAGACATTAATTTCCGCGTGG 657
Db 179 GlnghYserProvalAlalaglnalalaproalYalysghlnghlnghlnValaspilleProleu 198

```

```

QY 658 CGGATCTGCTCCCGACCCAGTTTGTGGCCATCTCGGAAAGGAGGCTTGACCAT 717
Db 199 ArgLeuLeuValProThrGlnTrpValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
718 AAGAATCTCATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisSerGlyGluAsnAlaGly 238
778 GCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 837
Db 239 AlaIleGlyLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
838 ATGATCTTGAATCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 897
Db 259 MetIleGlnIleMetHisLysGlnAlaLysAspTrpLysAlaAspGluValPro 278
898 CTGAATATCTTGGGACACATGCTGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
Db 279 LeuLysIleLeuAlaHisAsnAsnIleValGlyArgLeuIleGlyLysGluValArgAsn 298
958 TTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
Db 299 LeuLysLysValGluGlnAspThrGlnThrLysIleThrIleSerSerLeuGlnAspLeu 318
1018 AGCATATACACCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1077
Db 319 ThrLeuLysAspProGluGlnThrIleThrValLysGlyAlaIleGluAsnCysGlyAsn 338
1078 GCTGAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
Db 339 AlaGluGlnGlnIleMetLysLysValArgGluAlaTrpGluAsnAspValAlaIleMet 358
1138 AACACCCACTC-----GATACTTC----- 1158
QY 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValGlyLeuMetProAla 378
Db 1159 ---TTCAGGCTGTACCCCATCACCAG-----TTTGGCCGCTTCCCGCAT 1200
379 SerSerSerAlaValProProProSerSerValThrGlyAlaAlaProTrpGlySer 398
1201 CATCACTCTTATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 399 PheMetGlnAlaProGluGlnGlnMetValGlnValPheIleProIleGlnAlaValGly 418
1261 GCCATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 419 AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerTrpPheAlaSerAlaSer 438
1321 ATCAAGATGCCCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 439 IleLysIleAlaProProGlnTrpProAspSerLysValArgMetValIleThrGly 458
1381 CCACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 459 ProProGlnAlaIleGlnLysAlaGlnGlyArgIleLysLysLysGluGluAsn 478
1441 TTCTTTAACCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 479 PhePheGlyProLysGluGlnValLysLeuGlnThrHisIleArgValProAlaSerAla 498
1501 GCTGCGCGGGTATGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 499 AlaGlyArgValIleGlyLysGlyLysLysValLysAsnLeuGlnAsnLeuThrAla 518
1561 GCAGAGATCATGCTGCTCGTGAACCAACCCAGATGAAATAGAGAGAGAGAGAGAGAGAG 1620
Db 519 AlaGluValAlaValProArgAspGlnThrProAspLysLysAspGlnValIleValLys 538
1621 ATTATGGGAGATTTCTTGTCTAGCCAGACTGCAACGGGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 539 IleIleGlyHisPheTrpAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla 558

```

```

QY 1681 CAGGTGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1737
Db 559 GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnGlnAlaArgArgLys 577

RESULT 7
ID 088477 PRELIMINARY; PRT: 577 AA.
AC 088477;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Coding region determinant binding protein.
GN IGFBP1 OR CRDBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92217743; PubMed=1559612;
RA Bernstein P.L., Herrick D.J., Prokipcak R.D., Ross J.;
RT "Control of c-myc mRNA half-life in vitro by a protein capable of
RL binding to a coding region stability determinant.";
RN Genes Dev. 6:642-654(1992).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94158886; PubMed=8114742;
RA Herrick D.J., Ross J.;
RT "The half-life of c-myc mRNA in growing and serum-stimulated cells:
RL influence of the coding and 3' untranslated regions and role of
RN ribosome translocation.";
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94179348; PubMed=8132663;
RA Prokipcak R.D., Herrick D.J., Ross J.;
RT "Purification and properties of a protein that binds to the C-terminal
RL coding region of human c-myc mRNA.";
RN J. Biol. Chem. 269:9261-9269(1994).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97322234; PubMed=9178888;
RA Leeds P., Kren B.T., Boylan J.M., Betz N.A., Steer C.J.,
RL Gruppone P.A., Ross J.;
RT "Developmental regulation of CRD-BP, an RNA-binding protein that
RN stabilizes c-myc mRNA in vitro.";
[5]
RP SEQUENCE FROM N.A.
RX Ross J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A.,
RL Fleisig A.J.;
RT Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Embryo;
RL MEDLINE=2108560; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Stabli F., Suzuki R., Tomita M., Wagner U., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Grotstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni U., Maehima U., Nazarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald C., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseuki S.,

```

RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL; AF061569; AAC72743.1; -;
 DR EMBL; AK013940; BAB29071.1; -;
 DR HSSP; P11940; 1CVD.
 DR MGD; MGI:1890357; Igf2bp1.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR InterPro; IPR004087; KH dom.
 DR InterPro; IPR004088; KH type_1.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00013; KH; 2.
 DR Pfam; PF00076; xtm; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PS00084; KH_TYPE_1; 4.
 DR PROSITE; PS0102; RRM; 2.
 SQ SEQUENCE 577 AA; 63450 MW; EFBBIAP2FP9F0344 CRC64;

Alignment Scores:

Pred. No.: 1,13e-133 Length: 577
 Score: 1936.00 Matches: 381
 Percent Similarity: 79.10% Conservative: 77
 Best Local Similarity: 65.80% Mismatches: 95
 Query Match: 32.89% Indels: 26
 DB: 11 Gaps: 5

US-09-270-437d-8 (1-3283) x 088477 (1-577)

QY 73 ATGACAAAGCTTTATCATCGGGAACCTGACCCGCCGTCACCGCCGACGACCTCCGCGAG 132
 Db 1 MetAsnIleuTYrIleGIYAsnLeuAsnIleuSerValThrProAlaAspIleuGIuLys 20
 QY 133 CTCTTTGGGACACAGAGCTGCTCCGCGGACAGAGTCTGCTGAAGTCCGCGTACGCC 192
 Db 21 ValPheAlaGIuHISLysIleSerTySerGIYInPheLeuValIysSerGIYTyAla 40
 QY 193 TTGCTGACACTACCCGACCGACGACCTCGCGCCATCCGCGACGACCTCCGCGTAA 252
 Db 41 PheValAspCysProAspGIuHISLysIleValMetLysAlaIleGIuThrPheSerGIYLys 60
 QY 253 GTGGAATTCATGGGAAATCATGGAATGATTAATCACTTAATAAAGCTTAAGAGC 312
 Db 61 ValGIuLeuGIuGIuLysArgLeuGIuIleGIuHISerValProLysGIuLysSer 80
 QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCACTGACAGTGGAGGTGGATGACCT 372
 Db 81 ArgLysIleGIuIleArgAsnIleProGIuIleuArgTrpGIuValIleuAspSerLeu 100
 QY 373 TTGGCTCAATATGAGGACAGTGGAGATGGAACAATCAACATCAACACAGACAGAAACCGCC 432
 Db 101 LeuAlaGIuTYrGIuHISLysIleValGIuAsnCysGIuIleValAsnThrGIuSerGIuThrAla 120
 QY 433 GTTGTCAAGCTCATATGCAACAGAGAAAGCAAAATAGCCATGAGAGCTAAGC 492
 Db 121 ValValAsnValThrTySerAsnArgGIuGIuThrArgGIuAlaIleMetLysLeuAsn 140
 QY 493 GGGCATCAGTTTGAAGACTCTCTTCAAGATTCTCAATCCCGATGAAGAGGTGAGC 552
 Db 141 GYHISGIuLeuGIuHISLysIleValIleuLysValSerTyIleProAspGIuGIuIleThr 160
 QY 553 TCCCTTCGCGCCCTCAAGAGCCGAGCGTGGGAGCACTCTCCCGGAGCAAGCCAC 612
 Db 161 -----GlnGIYProGIuLeuGIYArgGIYGIYPheGIYSerArgGIYProArg 178
 QY 613 -----GCCCTGGGGGCACTTCTCAGGCGACAGATGATTTCCCGCTG 657
 Db 179 GlnGIYSerProValAlaIleGIYAlaProAlaLysGIuGIuProValAspIleProLeu 198
 QY 658 CGGATCTGTGTCACCCGAGTGTGTTGGTCCATCATCGGAAAGAGGGCTTGACATA 717
 Db 199 ArgLeuLeuValProThrGIuTYrValGIYAlaIleIleGIYLysGIuGIYAlaThrIle 218

QY 718 AAGACATCATTAACCAAGCCAGTCCCGGTAGATATTCATAGAAAAGAACTGTGA 777
 Db 219 ArgAsnIleThrIysGIuThrGlnSerIysIleAspValIleArgLysGIuHISLysIle 238
 QY 778 GCTGAGAGAGCCGTGACCATCCATGCCACCCGAGGGGACTCTTGACATGCCGC 837
 Db 239 AlaAlaGIuLysValIleSerValIleSerThrProGIuGIYCySerSerIleAspLys 258
 QY 838 ATGATTTCTTGAATCATGACAGAAAGGACAGATGAGCAACCTAGCCGAGATGCT 897
 Db 259 MetIleLeuGIuIleMetHISLysGIuAlaLysAspTrpLysThrAlaAspIleValPro 278
 QY 898 CTGAAATCTTGGCACACATGCTGCTGGTGAAGACTGATGGAAGAAAGACAGAAAT 957
 Db 279 LeuYsIleLeuAlaHISAsnAsnPheValGIYArgLeuIleGIYLysGIuLysArgAsn 298
 QY 958 TTGAAGAAATTTGAACATGAAACAGGACGCAAGATTAACATCTACTTTGACGAGATTG 1017
 Db 299 LeuYsLysValGIuGIuAspThrGIuThrIysIleThrIleSerSerLeuGIuAspLeu 318
 QY 1018 AGCATATTAACAACCCGAAAGAACCATCATCTGTGAAGGCGACAGTTGAGCCGTGCCAGT 1077
 Db 319 ThrLeuTYrAsnProGIuArgThrIleThrValIysGIYAlaIleGIuAsnCysArg 338
 QY 1078 GCTGAGATGAGATTATGAGAGAGCTGCGTGAAGCTTTGAAAATGATATGCTGCTGTT 1137
 Db 339 AlaGIuGIuGIuIleMetLysValArgGIuLysValArgGIuAsnAspValAlaAlaMet 358
 QY 1138 AACACCCACCTC-----GGATCTTC----- 1158
 Db 359 SerLeuGIuHISerHISLeuIleProGIYLeuAsnLeuAlaAlaValGIYLeuPheProAla 378
 QY 1159 ---TCCAGCTGTATCCCCCATCACAG-----TTGGCCGCTGCCCGAT 1200
 Db 379 SerSerSerAlaValProProProProSerSerValThrGIYAlaAlaProTySerSer 398
 QY 1201 CATCACTCTTATTCAGAGCAGAGAGATTGTGATCTCTTCATCCCAACCCAGCGCTGGGC 1260
 Db 399 PheMetGIuAlaProGIuGIuGIuMetValGIuValPheIleProAlaGIuAlaValGIY 418
 QY 1261 GCCATCATCGGAGAAAGAGGGGCGACACATCAACAGCTGGGAGATTGGCGGCTCT 1320
 Db 419 AlaIleIleGIYLysLysGIYGIuHISLysGIuIleuSerArgPheAlaSerAlaSer 438
 QY 1321 ATCAAGATTGCCCTCGGAGAGCCGACAGCTGACGCAAGAGATGATCATCATCACCGGG 1380
 Db 439 IleLysIleAlaProProGIuThrProAspSerLysValArgMetValIleThrGIY 458
 QY 1381 CCACCGGAGCCCATGTTCAAGGCCCAAGGACGATCTTTGGGAAACTGAAAGAGAAAC 1440
 Db 459 ProProGIuAlaGIuPheLysAlaGIuGIYArgIleTyGIYLysLeuLysGIuGIuAsn 478
 QY 1441 TTTCTTAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGATGCTCCCTCTCCACA 1500
 Db 479 PhePheGIYProLysGIuGIuValLysLeuGIuThrHISLysArgValProAlaSerAla 498
 QY 1501 GCTGCGCGGTGATTGGCAAGGTGGCAAGCGTGAACCACTGCGAAGCTTAACAGT 1560
 Db 499 AlaGIYArgValIleGIYLysGIYLysThrValAsnGIuLeuGIuAsnLeuThrAla 518
 QY 1561 GCAGAGTCACTGCTGCTGTGACCAACGCGCATGAAATGAGAAAGTATGTCGACGA 1620
 Db 519 AlaGIuValValValProArgAspGIuThrProAspGIuHISAspGIuValIleValLys 538
 QY 1621 ATTAATGGGAGCTCTTGTGACCGACATGCGACACCGCAAGATCAGGAAATTTGTACAA 1680
 Db 539 IleIleGIYHISpHeTYrAlaSerGIuMetAlaGIuArgLysIleArgAspIleLeuAla 558
 QY 1681 CAGGTAGACAGCAGAGAGAAATACCTCAGGAGAGTGGCTTCACAGCGCAGACAG 1737
 Db 559 GlnValLysGIuGIuHISLysGIuLysGIuSerHISLeuAlaGIuAlaArgArgLys 577

RESULT 8
Q9NZ18 PRELIMINARY; PRT; 577 AA.
AC Q9NZ18;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE mRNA-binding protein CRBP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M., Tsipalis C.M.,
RA Panoutarakopoulos G., Kyriazoglou I., Voutzoulis S., Tsipalis C.M.,
RA Kitis C., Agnantis N., Pandis N.;
RT "Ectopic expression of a KH-domain containing protein, highly
homologous to both human IMP-1 and mouse CRD-BP, in benign and
malignant mesenchymal tumors."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF198254; AAF37203.1; -.
DR HSBP; P1340; ICYJ.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rim; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63456 MW; 0749A060F252D81D CRC64;
Alignment Scores:
Pred. No.: 133e-133 Length: 577
Score: 1935.00 Matches: 382
Percent Similarity: 78.93% Conservative: 75
Best Local Similarity: 65.98% Mismatches: 96
Query Match: 32.87% Indels: 26
Gaps: 5
US-09-270-437d-8 (1-3283) x Q9NZ18 (1-577)
QY 73 ATGAAACAGCTTACATCGGAACTGAGCCCGCGTACCGCGACAGACTCCGGCAG 132
Db 1 MetAsnLysLeuYrILleglYAsnLeuAsnGluSerValThrProAlaAspLeuGluYs 20
QY 133 CTCTTTGGGACACGAAAGCTGCGCGGAGCAGGTCTCTGTAAGTCCGGCTACGCG 192
Db 21 ValPheAlaGluHnIshYsIleSerTYrSerGlyGlnPheValIysSerGlyTYrAla 40
QY 193 TTGTGAGACACCCCGACGAGCTGCGGCGCATCGAGCCCTCTCGGGTAA 252
Db 41 PheValAspCysProAspGlnHnIshTPAlaMetLysAlaIleGlnHnRhesSerGlyYs 60
QY 253 GTGGAATTGATGGGAAATCATGGAAGTTGATTACTCACTCTCTAAAAGCTAAGAGC 312
Db 61 ValGluLeuGlnIshYsArgLeuGlnIleGluHnIserValProIysIysGlnArgSer 80
QY 313 AGGAAATTCAGATTGGAACATCCCTCCATCGAGTGGAGGAGGTGTGATGACT 372
Db 81 ArgIlyIleGlnIshYsArgLeuGlnIleGluHnIserValProIysIysGlnArgSer 100
QY 373 TTGGCTCAATATGAGACAGTGAAGATGTGAAACAATGACACAGACAGAAACGCC 432
Db 101 LeuAlaGlnIshYrGlyThrValGlnAsnGluGlnValAsnThrIshYsSerGlnThrAla 120
QY 433 GTTGTCACCTCATATGCAACAGAGAGAGCAAAATAGCCATGAGAAAGCTAAGC 492
Db 121 ValValAsnValIshYrSerAsnArgGluGlnThrArgGlnAlaIleMetLysLeuAsn 140

QY 493 GGCATCAGTTTGAACACTCTCTTCAAGATTCTTACATCCGGATGAAGAGTGAC 552
Db 141 GlnIshGlnLeuGlnAsnHnIshAlaLeuYsValSerTYrIleProAspGlnGlnIleAla 160
QY 553 TCCCTTCGCCCCCTCAGCCGAGCCAGCTGGGACACACTTCCCGGACACAGGCCAC 612
Db 161 -----GlnGlyProGlnAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 613 -----GCCCTGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTG 657
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaIysGlnGlnGlnValAspIleProLeu 198
QY 658 CGGATTCCTGTGTCGCCACCCAGCTTGTGGTCCATCATCGAAGAGAGGCTTGACCTA 717
Db 199 ArgLeuLeuValProThrGlnTYrValGlyAlaIleIleGlyIysGlnGlyAlaThrIle 218
QY 718 AAGAAATCATCACTAAGCAGACCCAGTCCGGGTAGTATCCATAGAAAGAGAACTCTGA 777
Db 219 ArgAsnIleThrIysGlnThrGlnSerIshLeaPValHsArgLysGlnAsnAlaGly 238
QY 778 GCTGAGAGAGCCGTGACCATCATGCTCAGCCAGAGGGAGCTTGAAGCATGCGCG 837
Db 239 AlaAlaGlnIshYsAlaIleSerValHnIshSerThrProGlnGlyCysSerSerAlaCysIys 258
QY 838 ATGATTTCTTGAATATGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 897
Db 259 MetIleLeuGlnIleMetIshYsGlnIshYsGlnIshYsGlnIshYsGlnIshYsGlnIshYs 278
QY 898 CTGAAATCTTGGACACAAATGCGCTTGTGGAGACATGATGGAAGAGAGGAGGAGGAG 957
Db 279 LeuIshThrLeuAlaHsAsnAsnPheValGlyArgLeuIleGlyIysGlnGlyArgAsn 298
QY 958 TTGAAGAAATATGACATGAACAGAGGACAGATTAACATCTCTTGGAGATTTG 1017
Db 299 LeuIshYsValGlnGlnAspThrGlnHnIshIleThrIleSerSerLeuGlnAspLeu 318
QY 1018 AGCATATACACCCGAGAAAGAACATCACTGTGAAGGGGACAGCTTGAGGCTGTGCCAGT 1077
Db 319 ThrLeuTYrAsnProGlnIshYrThrIleThrValIshYsGlyAlaIleGlnIshYsCysArg 338
QY 1078 GCTGAGATGAGATTTATGAGAAAGCTGCGGTGAGGCTTGAATATGATGCTGGTGT 1137
Db 339 AlaGlnGlnIshYrIleMetLysValArgGlnAlaIshYrGlnIshYsValAlaAlaIleMet 358
QY 1138 AACACCCACTCC-----GATATCTC----- 1158
Db 359 SerLeuGlnSerHnIshLeuThrProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
QY 1159 ---TCCAGCTGTACCCCATCAG-----TTGGCCCGCTCCCGCAT 1200
Db 379 SerSerSerAlaValProProProProSerSerValThrGlyAlaAlaProIysSerSer 398
QY 1201 CATCACTTTATCCAGAGAGAGATGTGAATCTCTTCATCCCAACCGAGCTGTGGC 1260
Db 399 PheMetGlnAlaProGlnGlnGlnMetValGlnValPheIleProAlaGlnAlaValGly 418
QY 1261 GCCATCATCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 419 AlaIleIleGlyIshYsGlyGlnHnIshIleYsGlnIshYsSerArgPheAlaSerAlaSer 438
QY 1321 ATCAAGATTGCCCTGCGGAGAGCCGAGAGCTGACGAAAGATGTGATCATCACCGG 1380
Db 439 IleIysIleAlaProProGlnIshYrProAspSerIshYsValArgMetValIleIleHnGly 458
QY 1381 CGACCGGAGAGCCAGTTCAAGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 459 ProProGlnIshYsGlnIshYsAlaGlnGlyArgIleYrGlyIshYsLeuIshYsGlnIshYs 478
QY 1441 TTCTTTAACCCCAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Db 479 PhePheGlyProIysGlnGlnValIshYsLeuGlnIshYsIleArgValProAlaSerAla 498

QY 1501 GCTGCGCGGATGTTGGCAAGGTGGCAAGACCGTGAACGAACCTGCAAGAACTTAAACAGT 1560
DB 499 AAGYArGValIleelYlYsGlyGlySThrValAsnGluLeuGlnLeuThraIa 518
QY 1561 GCAGAAGTCATGCTGCTGCTGACCAAGCCAGATGAAAGAAATGAGAAAGTATGCTGACA 1620
DB 519 AAGlValValValProArgAspGlnThrProAspGluAsnAspGlnValIleVallys 538
QY 1621 ATTATCGGCACTCTTCTTGTCTGACCGACTGCAAGCGCAGAGTCAAGAAATTGTACAA 1680
DB 539 IleIleGlyHisPheTyrIleAserGlnMetAlaGlnAspGlyIleArgAspIleLeuAla 558
QY 1681 CAGGTGAAGCAGCAGAGCAGCAAAATACCTTCAGGAGTGCCTCAGACAGCAGCAAG 1737
DB 559 GlnVallyslGlnGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAlaArgArglys 577
RESULT 9
Q8BRH1
ID Q8BRH1 PRELIMINARY; PRT; 577 AA.
AC Q8BRH1;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Inulin-like growth factor 2.
GN IGFBP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK044850; BAC32119.1; -.
DR MGD; MGI:1890357; Igfbp1.
DR GO; GO:0003676; P.nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type_1.
DR InterPro; IPR005054; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rim; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RMW; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RM; 2.
SQ SEQUENCE 577 AA; 63378 MM; D75897450841532E CRC64;
Alignment Scores:
Pred. No.: 3.69e-133 Length: 577
Score: 1929.00 Matches: 380
Percent Similarity: 78.93% Conservative: 77
Best Local Similarity: 65.63% Mismatches: 96
Query Match: 32.77% Indels: 26
DB: 11 Gaps: 5
US-09-270-437D-8 (1-3283) x Q8BRH1 (1-577)
QY 73 ATGAACAAGCTTATACCTGGGAACCTGAGCCGCGCTCAACCCGACAGCACTTCGGGAC 132
DB 1 MetAsnIleuIleuIleGlyAsnIleuAsnGlnSerValThrProAlaAspLeuGlnLys 20
QY 133 CTCCTTGGGAGCAGAGCTGCCCTGGCGGACAGAGTCTCTGTAAGTCGGGCTACGCC 192
DB 21 ValPheAlaGlnHisIleSerTyrSerGlyGlnPheIleuVallysserGlyTyrAla 40
QY 193 TTGCTGACATACCCCAAGCAAGAACTGGGCACTCCGCCCATGACAGACCTCTCGGGTAA 252
DB 41 PheValAspCysProAspGlnHisThrPalaMetLysAlaIleGlnThrPheSerGlyLys 60

QY 253 GTGAATTCATGGGAAATATCATGGAAGTTGATTAATCTCAATCTTAAAGGCTAAGGAC 312
DB 61 ValGlnLeuGlnGlyLysArgLeuGlnIleGlnHisSerValProLysGlnArgSer 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCCATCCGAGCTGGGAGAGTGTGATGACACT 372
DB 81 ArgLysIleGlnIleArgAsnIleProProGlnLeuAspGlnValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGAGCAGATGAGAGATGTGAAACAAGTCAACACAGACACAGAAACGCC 432
DB 101 LeuAlaGlnTyrGlyThrValGlnAsnGlyGlnGlnValAsnThrGlnSerGlnThrAla 120
QY 433 GTTTCACATGTCATATGCAACAAGAGAGAGCAAAATAGCCATGAGAGAGTAAAC 492
DB 121 ValValAsnValThrTyrSerAsnArgGlnGlnThrArgGlnAlaIleMetLysLeuAsn 140
QY 493 GGGCATCAGTTTGAAGAACTACTCCCTCAAGATTTCTCATTCGCCGATGAGAGGTAGC 552
DB 141 GlyHisGlnLeuGlnAsnHisAlaLeuLysValSerTyrIleProAspGlnGlnIleThr 160
QY 553 TCCCTTCGCCCTCAGACGAGCCAGCGTGAGGACCACTCTCCCGGAGCAGAGCCAC 612
DB 161 -----GlnGlyProGlnAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 613 -----GCCCTGGGGGACCTTCTCAGGCGCAGACAGATGATTTCCCGCTG 657
DB 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
QY 658 CGGATCCGTCGCGCCACCCAGTTTGTGTCCTCATTCGGAAGAGAGGCTTACCATTA 717
DB 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGlnGlyAlaThrIle 218
QY 718 AAGAATCATCATGACAGACCCAGTCCCGGATAGATATCCATGAAAGAAAGAACTCTGCA 777
DB 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGlnAsnAlaGly 238
QY 778 GCTCAGAGAAAGCTGTTCACCATTCATCCACCCAGAGGGGACTTGAAGCATGCCGC 837
DB 239 AlaAlaGlnLysAlaIleSerValHisSerThrProGlnGlyCysSerSerAlaCysLys 258
QY 838 ATGATTTCTGAATATCATGACAGAAAGAGCAGATGACCAAACTAGCCGAAGAACTTCT 897
DB 259 MetLeuGlnIleMetHisLysGlnAlaLysAspThrLysThrAlaAspGlnValPro 278
QY 898 CTGAATATCTTGACACCAATGAGCTGTGTAAGATGTAATGGAAGAAGCAGAAAT 957
DB 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlnGlyArgAsn 298
QY 958 TTGAAGAAAATTGAACATGAAACAGGACCAAGATTAACAATCTCATCTTTGACAGATTGG 1017
DB 299 LeuLysLysValGlnGlnAspThrGlnThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 1018 AGCATATACAAACCCGAAAGAAACCATCATCTGTGAAGGACAGTTGAGGCTTGCCAGT 1077
DB 319 ThrLeuTyrAsnProGlnArgThrIleThrValLysGlyAlaIleGlnLeuAsnGlyCysArg 338
QY 1078 GCTGAGATAGAGATTGAAGAGCTGCGTGAAGGCTTGAATAATGATATGCTGCTGT 1137
DB 339 AlaGlnGlnGlnIleMetLysValArgGlnAlaTyrGlnAsnAspValAlaAlaMet 358
QY 1138 AACACCACTCC-----GGATCACTTC----- 1158
DB 359 SerLeuGlnSerHisIleuIleProGlyLeuAsnLeuAlaValGlyLeuPheProAla 378
QY 1159 ---TCCAGCGCTGTAACCCCATCAGCAG-----TTGGCCCGCTTCCGCAT 1200
DB 379 SerSerSerAlaValProProProProSerSerValThrGlyAlaAlaAlaProTyrSerSer 398
QY 1201 CATCACTTATACAGAGCAGAGATTTGAAATCTTCAATCCCAACCCGATGCTGGG 1260
DB 399 PheMetGlnAlaProGlnGlnGlnMetValGlnValPheIleProAlaGlnAlaValGly 418

Db	239	AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys	258
QY	838	ATGATTTCTTGAATTCATGACGAAAGCGCAGATGAGACCAATCCAGCAAGATTCCT	897
Db	259	MetIleLeuGluIuIleMetHisLysGluAlaLysAspThrLysThrAlaAspGlyValPro	278
QY	898	CTGAAAAATCTTGGCACAACAATGGCTTGTGGGAAGCTGATTTGGAAAAAGCGAGAAAT	957
Db	279	LeuLysIleLeuAlaHisAsnAspHeValGlyArgLeuIleGlyGluGlyArgAsn	298
QY	958	TTGAAGAAATTTGAACATGAAACAGGGACCAAGATACAAATCTCATCTTTGGACGATTTG	1011
Db	299	LeuLysLysValGluGlnAspIlnrGluThrLysIleThrLisSerSerLeuGlnAspLeu	318
QY	1018	AGCATATACACCCCGGAAAGAACCATCACTGTGGAAGGGCAGACAGTTGAGGCGCTGTCCAGT	1077
Db	319	ThrLeuTyAsnProGluIuArgThrIleThrValLysGlyAlaIleGluAsnCySerArg	338
QY	1078	GCTGAGATGAGATTATGAAAGACTGCTGAGCGCTTGTGAATAATGATATGCTGTGCTGT	1137
Db	339	AlaGluGlnGluIleMetLysLysValArgGluAlaTyGluAsnAspAlaAlaIleMet	358
QY	1138	AACACCCATCC-----GGATCTTC-----	1158
Db	359	SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuAspHeProAla	378
QY	1159	---TCCAGCCTGTACCCCCATCACAG-----TTTGGCCCGGTCCCGCAT	1200
Db	379	SerSerSerAlaValProProProProSerSerValThrGlyAlaAlaProTySerSer	398
QY	1201	CATCACTCTTATCCAGACGAGAGATTGTGAATCTTTCATCCCAACCAGCGTGTGGC	1260
Db	399	PheMetGlnAlaProGluGlnIleMetValGlnValPheIleProAlaGlnAlaValGly	418
QY	1261	GGCATCATCGGGAGAGAGGGGGCCACACATCAAACAGCTGGCGAGATTGGCCGAGCTCT	1320
Db	419	AlaIleIleGlyLysLysGlyGlnHisIleLysGlnLeuSerArgPheAlaSerAlaSer	438
QY	1321	ATCAAGATTGCCCTCGCGGAGAGGCCCAACATCGACAGCAAGATGTGTCATCATCACCGG	1380
Db	439	IleLysIleAlaProProGluIuThrProAspSerLysValArgMetValValIleThrLys	458
QY	1381	CCACCGGAAACCCAGTTCAAGGCCCAAGGACGATCTTTGGGAAACTGAAAGAGAAAC	1440
Db	459	ProProGluAlaGlnPheLysAlaGlnGlyArgIleTyGlyLysLeuLysGlnGluAsn	478
QY	1441	TTCTTTAACCCCAAGAGAAAGTGAAGCTGGAAGGGCATATTCAGAGTCCCTCTTCCACA	1500
Db	479	PhePheGlyProLysGluGlnValLysLeuGlnThrHisIleArgValProAlaSerAla	498
QY	1501	GCTGGCCGGGTGATTTGGCAAGAGTGGCAAGACCGTGAACGAATCTGAGAAATTGAACAGT	1560
Db	499	AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnSerThrAla	518
QY	1561	GCAAGATCATGCTGCTCTGTGACCAACGCCAGATGAAATGAGAAAGTATGTCGACGA	1620
Db	519	AlaGluValAlaValProArgAspIlnrThrProAspGluAsnAspGlnValIleValLys	538
QY	1621	ATTATCGGGCACTTGTGTATGACCAAGCTCAACAGCGCAAGATCAGGAAATTTGTACA	1680
Db	539	IleIleGlyHisPheTyAlaSerGlnMetAlaGlnArgLysIleArgAspIleLeuAla	558
QY	1681	CAGGTGAAGCAGCAGGACAGAAATTAACCTCAGGAGAGTCCGCTCACAGCGAGACAG	1737
Db	559	GlnValLysGlnGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAlaArgArgLys	577
RESULT 11			
Q9PM80		PRELIMINARY, PRT, 582 AA.	
AC	Q9PM80;		
DT	01-MAY-2000 (TremBLrel. 13, Created)		
DT	01-MAY-2000 (TremBLrel. 13, Last sequence update)		
DT	01-OCT-2003 (TremBLrel. 25, Last annotation update)		

DE Vg1 RNA binding protein (Decapentaplegic and Vg-related 1, RNA binding protein).

GN DVIRBP.

OS *Brachydanio rerio* (zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; OC Cyprinidae; Danio.

NCBI_taxid=7955;

RN [1]

RP SEQUENCE FROM N.A.

RA Zhang O., Yaniv K., Oberman F., Wolke U., Git A., Fromer M., RT Taylor W., Meyer D., Standart N., Raz E., Yisraeli J.K.;

RT "Vg1 RBP intracellular distribution and evolutionarily conserved expression suggest multiple roles during development.";

RL Mech. Dev. 0:0-0(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=AB; TISSUE=Body;

RX MEDLINE=22386257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., RA Klausner R.D., Collins F.S., Wanner I., Shenmen C.M., Schler G.D., RA Altchul S.F., Zeeberg B., Bietow K.H., Schaefer C.F., Bat N.K., RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E., RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C., RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., RA Bosak S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunatille P.H., RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., RA Fahey U., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., RA Kryzhaniski M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E., RA Jones S.J., Maira M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=AB; TISSUE=Body;

RA Strausberg R.;

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.

DR EMBL: AP161270; AAD45610.1; -

DR EMBL: BC045873; AAR45873.1; -

DR ZFIN: ZDB-GENE-000308-1; dvrltbp.

DR GO: GO:0003676; Functional acid binding; IEA.

DR InterPro: IPR004087; KH dom.

DR InterPro: IPR004088; KH type 1.

DR InterPro: IPR000504; RNA_rec_mot.

DR Pfam: PF00013; KH; 4.

DR Pfam: PF00076; Trm; 2.

DR SMART: SM00322; KH; 4.

DR SMART: SM00360; RRM; 2.

DR PROSITE: PSS0084; KH TYPE_1; 4.

DR PROSITE: PSS0102; RRM; 2

SO SEQUENCE 582 AA; 63351 MW; 9DAE63200681B306 CRC64;

Alignment Scores:

Pred. No.:	2,52e-130	Length:	582
Score:	1890.50	Matches:	377
Percent Similarity:	78.63%	Conservative:	72
Best Local Similarity:	66.02%	Mismatch:	97
Query Match:	32.12%	Indels:	25
DB:	13	Gaps:	5

US-09-270-437D-8 (1-3283) x Q9PM80 (1-582)

QY 133 CTCCTTTGGGAGACAGAAAGCTGCCCTGGGAGACAGCTCCTGCTGAAGTCCGGCTAGCC 192
 Db 21 lIepheluglntrpIysIleProhPheSerAlaProhPheLeuValIySergIlyTala 40
 QY 193 TTCGTGACATACCCCGACCAACATCGGGCCATCGCGCATCGAGACCTCTCGGGTAA 252
 Db 41 PheValAspCysProAspGluValAlaMetArgAlaIleAspThrLeuSergIlyS 60
 QY 253 GTGGAAATTCAGTGGGAAAATCATGAGATGATTACTGACTGCTTAAAGCTAAGAGC 312
 Db 61 ValGlnLeuHISglYlyValIleuglIyValGlnHISerValProlySArgGlnArSerg 80
 QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCAACCTCGAGCTGGGGGTGGTGGAGACTT 372
 Db 81 CysIysLeuGlnIleArgAsnIleProhPheHISmetGlntrpGlnValLeuSergIlyeu 100
 QY 373 TTGGCTCAATATGGGACAGTGAAGATGTGAACAAGTCAACACAGACAGAAACCCGC 432
 Db 101 LeuAlaGlntrpGlyThrValGlnSerCysgluGlnValAsnThrAspThrGlnThrAla 120
 QY 433 GTTGTCACAGCTCATATATGCAACAGAGAGAGAGAAAATAGCCATGAGAGAGTAAAC 492
 Db 121 ValValAsnValArgTyGlyAlaIyAspGlnAlaArgGlnAlaMetAspIySLeuAsn 140
 QY 493 GGGCATCAGTTTGAGACTACTCCTTCAAGATTTCTCATCCCGGATGAGAGAGGTGAGC 552
 Db 141 GlyPheLeuMetGlnAsnTrpAlaLeuIySValSeryIlyleProAspGlnThrAlaIa 160
 QY 553 TCCCTTGGCCCTCAGCAGAGCCAGCGTGGG-----GACCACTCTTCCGGAGCA 606
 Db 161 AlaAspAlaProAlaValGlyGlyArgArgGlyPheAsnProArgGlyProArgGln 180
 QY 607 GGGCAGCCCTGGGGGACATTCTCAGGCGACAGATGATTTCCCGTGGCATCTCTG 666
 Db 181 GlySerProSerLeuGlyAlaArgProIySLeuGlnSerAspValProhLeuArgLeu 200
 QY 667 GTCCCAACCCAGTTGTTGTCGCCATCATCGGAAGAGAGGGCTTACCACTAAGAACATC 726
 Db 201 ValProIyGlnPheValGlyAlaIleIleGlyIySglnGlyAlaThrIleArgAsnIle 220
 QY 727 ACTAGACAGACCCAGCCCGGAGATATCATAGAAAAGAAACTCTGAGCTGACAG 786
 Db 221 ThrIySglnThrHISerIySleAspIleHISArgIyGlnAsnAlaGlyAlaIleGln 240
 QY 787 AAGCTGTACCATCATGATCCACCCAGAGGGGACTTTCAGAGCATGCCGATGATTTCT 846
 Db 241 IySProIleThrValHISerThrProIyGlnIyCysSerSerAlaCysArgAsnIleMet 260
 QY 847 GAAATATGACAGAAAGACAGATGAGACAACTAGCCGAGAGATTTCTCGAATAATC 906
 Db 261 GlnIleMetGlnIySglnAlaIleAspThrIySleThrGlnIyIleProLeuIySle 280
 QY 907 TTGGACACACATGGCTGTGGTGAAGACTGATTTGAAAAGAGGCAAAATTTGAAGAA 966
 Db 281 LeuAlaHISAsnAsnPheValGlyArgLeuIleGlyIySglnIyArgAsnLeuIyS 300
 QY 967 ATTGAACATGAAAGAGGACCAAGATPAACATCTCTTTGGAGATTTGACATATAC 1026
 Db 301 IleGlnIyAspThrAspThrIySleThrIleSerProLeuGlnAspLeuThrLeuIy 320
 QY 1027 AACCCGAGAAAGAACCTCATCTGTGAAGGGAGACAGTGGCCGTGCGATGCTGAGAA 1086
 Db 321 AsnProGlnArgThrIleThrValIySglyThrLeuAspAlaCysAlaIySArgIyGln 340
 QY 1087 GAGATTATGAAAGAGCTGCGTGGAGCCCTTGAATAATGATATGCTGCTGTAAACCCAC 1146
 Db 341 GlnIleMetIySValArgGlnSerIyGlnAsnAspAlaAlaIleMetHISLeuGln 360
 QY 1147 TCC-----GGATCTTC----- 1158
 Db 361 SerAsnLeuIleProGlyLeuAsnLeuAlaIleugIyLeuPheProGlyAlaAlaSer 380

QY 1159 TCCAGCCTGTAACCCCATCACACAGTTTGCCCGTTCG-----CATCATCAC 1206
 Db 381 GlyIyIleSerProSerValIalSergIlyProProGlyAlaGlnAlaGlyTyrGln 400
 QY 1207 TCTTAT-----CCAGACAGAGATTTGTGATCTCTTCACTCCCAACCAAGGCTGTG 1257
 Db 401 SerPheGlyAlaGlnMetGlnSerGlnThrValHISleuPheIleProAlaLeuAlaVal 420
 QY 1258 GGGCCATCATCGGAGAAAGGGGACACATCAACACAGTGGGAGATTCGCGGAGCC 1317
 Db 421 GlyAlaIleIleGlyIySglnGlyGlnHISleIySglnLeuSerArgPheAlaGlyAla 440
 QY 1318 TCTATCAAGATTTGCTCGGAGAGCCAGACGTCACCGAAAGATGCTCATATCAC 1377
 Db 441 SerIleIySleAlaProAlaAspGlyIleAspAlaIySArgMetValIleIleSer 460
 QY 1378 GGGCCACCGGAGAGCCGATTCAGGCGCAGAGGAGCGATCTTGGGAACTGAAAGAGAA 1437
 Db 461 GlyProGlnAlaGlnPheIySAlaGlnIyArgIlePheGlyIySLeuIySglnGln 480
 QY 1438 AACTCTTTAACCCCAAGAAAGATGAAAGCTGGAAGCGCATATCAGATGCTCTTCC 1497
 Db 481 AsnPhePheGlyProIySglnGlnValIySLeuGlnAlaHISleIyValProSerPhe 500
 QY 1498 ACAGCTGCGCGGTGATTTGCAAAAGTGGCAAGCCGTGAACGAACTGCAAGAACTTAAC 1557
 Db 501 AlaIleGlyArgValIleGlyIySglyGlyIySThrValAsnGlnLeuGlnAsnLeuThr 520
 QY 1558 AGTGAGAAAGTCATCGTGCCTCGTACCAACGCGCAGATGAAATGAGGAATGATCTGC 1617
 Db 521 SerAlaGlnValAlaValProArgAspGlnThrProAspGlnAsnAspGlnAlaVal 540
 QY 1618 AGAATTATGCGGCACTTCTTGTAGCCAGACTGCAACGCGCAAGATCGAGAAATGTGA 1677
 Db 541 IySleIleThrGlyHISPheIyAlaSerGlnLeuAlaGlnArgIyIleGlnIyIleIle 560
 QY 1678 CAACAGGTGAGACAGACAGAGAGCAAAATACCT 1710
 Db 561 SerGlnValArgArgGlnGlnIleProIySPro 571
 RESULT 12
 QY 08C2J9 PRELIMINARY; PRT; 579 AA.
 AC 08C2J9;
 DT 01-MAR-2003 (TRMBLrel. 23, Created)
 DT 01-MAR-2003 (TRMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TRMBLrel. 25, Last annotation update)
 DE Insulin-like growth factor 2.
 GN IGF2BP3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=MOD; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK08465; BAC40370.1; -.
 DR MGD; MGI:1890359; IGF2bp3.
 DR GO; GO:0003676; F.nucleic acid binding; IPA.
 DR InterPro; IPR004087; KH dom.
 DR InterPro; IPR004088; KH_type_1.
 DR InterPro; IPR000504; RNA_rec_moc.
 DR Pfam; PF00013; KH; 4.
 DR Pfam; PF00076; rrm; 2.
 DR SMART; SM00322; KH; 4.
 DR SMART; SM00360; RRM; 2.
 DR PROSITE; PSS0084; KH_type_1; 4.

DR PROSITE: PSS0102; RM: 2;
SQ SEQUENCE 579 AA; 63551 MW; 937E601A95D06B77 CRC64;

Alignment Scores:

Pred. No.: 3, 47e-129 Length: 579
Score: 1875.00 Matches: 372
Percent Similarity: 78.22% Conservative: 77
Best Local Similarity: 64.81% Mismatches: 87
Query Match: 31.86% Indels: 38
DB: 11 Gaps: 6

US-09-270-437d-8 (1-3283) x Q8C209 (1-579)

```

QY 73 ATGAACAAGCTTTATCATCGGGAACCTGAGCCCGCCGTCACCGCCGAGCACTCCGAGC 132
Db 1 MetLeuLysLeuLysIleGlyAsnLeuSerAspHisAlaGlyProAlaAspLeuGlySer 20
QY 133 CTCCTTTGGGACAGAGAACCTGCGCCGCGAGACGTCCTGCTGAAGTCCGCTACGCC 192
Db 21 ValPheLysAspAlaLysIleProValAlaGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TTGCGAGCTACCCCGACAGAACCTGCGCCATCCGCGCATCGAGACCTCTCGGTTAA 252
Db 41 PheValAspLysProAspGlyGlyTyrPheLeuLysAlaIleGlyAlaLeuSerGlyLys 60
QY 253 GTGGAATTCAGTGGGAAATCATGGAAGTTGATTACTACGTCTTAAAGCTAAGAGAC 312
Db 61 MetGluLeuHisGlyLysProMetGluValGluHisSerValProLysAlaGlnArgIle 80
QY 313 AGGAAATTCAGATTCCGAACATCCCTCTCAGCTGACAGTGGAGAGTGTGGATGACTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnIleProGluValLeuAspSerLeu 100
QY 373 TTGCTCATATGAGGAGAGATGAGATGGAACATGTCACACAGACAGAAACGGCC 432
Db 101 LeuValGlnIleTyrGlyValValGlnSerCysGlnIleValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAAGCTCATATGCAACAAGAGAAAGCAAAATATAGCCATCGAGAGACTAAGC 492
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATATGTTTGAACATACTCTTCAGATTTCTTACATCCCGGATGAAAGAGTGAAC 552
Db 141 GlyPheGlnLeuGlnAsnPheThrLeuLysValAlaIleThrLeuProAspGluThrAla 160
QY 553 -----TCCCTTGGCCCTCCAGCGAGCCAGCGTGGGAC-----CACTTTC 597
Db 161 GlnGlnAsnProSerProGlnLeuArgGlyArgArgGlyProGlnArgGlySerSer 180
QY 598 CGGAGAGCAAGCGCCCTCGGGGAGCACTTCTCAGCGCCAGACAGATTTATTCCTCGCG 657
Db 181 ArgGlnAla-----SerProGlySerValSerLysGlnLysProCysAspLeuProLeu 198
QY 658 CGGATCTGTGTCACCCACCCAGTTTGTGGCATCATCGGAAAGAGGAGCTTGACCATTA 717
Db 199 ArgLeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGlnLysGluAlaThrIle 218
QY 718 AAGAAATCATCTAAGAGACCCAGTCCCGGTGATATTCATAGAAAGAACTCTGGA 777
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGlnAsnThrGly 238
QY 778 GCTGCAGAGAGCCTGTACATCCATCCAGCCCGAGAGGGGCTTGTGAAGATCGCGC 837
Db 239 AlaAlaGlnLysSerIleThrIleLeuSerThrProGlnLysThrSerAlaAlaCysLys 258
QY 838 ATGATTTCTGAATCATGAGAAAGAGAGATGAGACCAATAGCCGAGAGATTTCTCT 897
Db 259 SerIleLeuGlnIleMetHisLysGlnAlaGlnAspIleLysPheThrGlnGluIlePro 278
QY 898 CTGGAATCTTGCCACACATAGCTGTTGTTGAAGCTGATTTGAAAGAGAGCAGAAAT 957
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlnGlyArgAsn 298

```

```

QY 958 TTGAAGAAATTTGAACATGAAACAGGACCAAGATTAACATCTCTTTGAGAGATTGG 1017
Db 299 LeuLysLysIleGlnGlnAspThrAspThrLysIleThrIleSerProLeuGlnIleu 318
QY 1018 AGCATATTAACCCCGAAAGAAACCATCATCTGTGAAGGACAGTGTGAGCCGTGCCACT 1077
Db 319 ThrLeuTyrAsnProGlnAlaArgThrIleThrValLysGlySerValGlnThrLysAlaLys 338
QY 1078 GCTGAGATGAGATTATGAGAAAGCTGCGTGAAGGCTTGTGAAGATGATGATGATGATG 1137
Db 339 AlaGlnGlnGluIleMetLysLysIleArgGlnSerTyrGlnAsnAspIleAlaSerMet 358
QY 1138 AAGACCCAC----- 1146
Db 359 AsnLeuGlnAlaAsnLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProPro 378
QY 1147 -----TCCGATTAATCTTCACCCCTGTACCCCATACCAAG 1182
Db 379 ThrSerGlyMetProProProThrSerGlyProProSerThrLeuThrPro----- 395
QY 1183 TTTGCGCCGTTCCCGCATCATCACTTTATCCAGAGAGAGATTTGATCTTTGATC 1242
Db 396 -----ProTyrProGlnPheGlnGln-----SerGlnThrGlnThrValHisLeuPheIle 412
QY 1243 CCAACCCAGAGCTGTGGCGCCATCATGTGGGAAAGAGGAGGACACATCAAAACGCTGAGC 1302
Db 413 ProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSer 432
QY 1303 AGATTGCGCGGAGCTCTATCAAGATTTGCCCTGCGGAAAGGCCAGACGTCAGCGAAAG 1362
Db 433 ArgPheAlaGlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArg 452
QY 1363 ATGTGTATCATACCGGCGCCACCGAAGCCAGTTCAGAGCCGAGAGAGATCTTTGGG 1422
Db 453 MetValIleIleThrGlyProProGlnAlaGlnPheLysAlaGlnGlyArgIleTyrGly 472
QY 1423 AAAGTGAAGAGGAAATCTTTAACCCCAAGAGAGAGTGAAGCTGGAAGCCCATATC 1482
Db 473 LysIleLysGlnGlnAsnPheValSerProLysGlnGluValLysLeuGlnAlaHisIle 492
QY 1483 AGAGTGGCTCTTCCACAGCTGCGCGGTGATTTGGCAAGAGTGGCAAGCCGTGAACGAA 1542
Db 493 ArgValProSerThrPheAlaAlaGlyArgValIleGlyLysGlyLysLysThrValAsnGln 512
QY 1543 CTGCAGAACTTAACAGTGCAGAGATCATGCTGCTGTCGACCAAGCCAGATGAAAT 1602
Db 513 LeuGlnSerLeuSerSerAlaGluValAlaValProArgAspGlnThrProAspGluAsn 532
QY 1603 GAGAAAGTGAATGCTCAGAATTTATCGGCACTTCTTGTGTAAGCCAGACTGCAAGCGCAG 1662
Db 533 AspGlnValAlaValLysIleThrGlnHisPheTyrAlaCysGlnValAlaGlnArgLys 552
QY 1663 ATCAGGGAATTTGTACAAAGCTGGAAGCCAGAGCAAGCAAGAA 1704
Db 553 IleGlnGlnIleLeuThrGlnValLysGlnHisGlnGln 566

```

RESULT 13

```

O9CPN8
ID O9CPN8; PRELIMINARY; PRT; 579 AA.
AC O9CPN8;
DT 01-UTN-2001 (TREMBLrel. 17, Created)
DT 01-UTN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE 10 days embryo cDNA, RIKEN full-length enriched library,
DE clone:2610036818, full insert sequence (igf2 mRNA-binding protein 3)
DE (Insulin-like growth factor 2, binding protein 3).
GN IGF2BP3 OR 261010N1IKIT OR MIMP3.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Atzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guernicich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Matzarelli U., Mombere P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Teyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.",
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,
 RA Yusa Y., Takeda M., Okano H.,
 RT "Expression of mouse 1912 mRNA-binding protein 3 and its implications
 for the developing central nervous system.",
 RL J. Neurosci. Res. 0:0-0(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain, and Olfactory epithelium;
 RX MEDLINE=22388257; Pubmed=12477932;
 RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Schemm C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Locantello N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
 RA Bosak S.A., McKeown P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy U., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Matra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain, and Olfactory epithelium;
 RA Strussberg R.,
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
 DR EMBL: AK011689; BAB27779.1; -;
 DR EMBL: AB046173; BAB19755.1; -;
 DR EMBL: BC045138; AAH45138.1; -;
 DR EMBL: BC049082; AAH49082.1; -;
 DR MGI:1890359; Igt2b3.
 DR GO:0003676; F: nucleic acid binding; IEA.
 DR InterPro: IPR004087; KH dom.
 DR InterPro: IPR004088; KH type_1.
 DR InterPro: IPR00504; RNA_rec_mot.
 DR Pfam: PF00013; KH; 4.
 DR Pfam: PF00076; trm; 2.
 DR SMART: SMO0322; KH; 4.
 DR SMART: SMO0360; RRM; 2.
 DR PROSITE: PSS0084; KH_type_1; 4.

DR PROSITE: PSS0102; RRM; 2.
 SQ SEQUENCE 579 AA; 63574 MW; CABD9A4355B392B7 CRC64;
 Alignment Scores:
 Pred. No.: 3,47e-129 Length: 579
 Score: 1875.00 Matches: 372
 Percent Similarity: 78.22% Conservative: 77
 Best Local Similarity: 64.81% Mismatches: 87
 Query Match: 31.86% Indels: 38
 DB: Gaps: 6
 US-09-270-437d-8 (1-3283) x Q9C9N8 (1-579)
 QY 73 ATGAACAAGCTTTACATCGGAGACCTTGAGCCCGCCGACCGGACGACCTCCGGGAC 132
 Db 1 MetanyleuylYrillleGlyAsnleuSeraspHisalaglProvalAaspneuIuser 20
 QY 133 CTCTTTGGGAGCAGAGAGCTCCCGCCGGGAGCAGGCTGCTGAATCCGGCTACGCC 192
 Db 21 ValPheylAspAlalylleProvalAlaglProPheleuValylsthrGlylYrila 40
 QY 193 TTGGTGACCTACCCGACCAAGACTGGGCGCATCCGCCATCGAGACCCCTCGGGTAA 252
 Db 41 PheValaspCyproAepgluglyTrrAlaleuylsalalileglAlaleuSerGlyls 60
 QY 253 GTGGAATTCATGCGGAAATCATGGAAGTTGATTCTAGTTCTTAAAGCTAAGAGC 312
 Db 61 MetGluLeuHlsglylYpProMetGluValGluHlsserValProlylArgGlnArglle 80
 QY 313 AGGAAATTCATGATTCGGAAGAACATCCCTCCTACCCGCGAGTGGGAGGTTGGATGACTT 372
 Db 81 ArgylsleuGlnlleArgasnlleProProHlssleuGlnlrrglValleuAspserleu 100
 QY 373 TTGGCTCAATGAGGAGAGTGGAGATGAGAACAGTCAACAGACAGACAGAAACCGCC 432
 Db 101 LeuValGlnlYrGlyValValGluSerCysgluGlnValasnthrAspserGlnhrAla 120
 QY 433 GTTGTCACATGACATCATGATGCAACAAGAGAGCAAGAAATGCGCATGGAAGAGTAAAC 492
 Db 121 ValValasnlValThrlylSerSerlylspGlnAlaArgGlnAlaleuAspysleuasn 140
 QY 493 GGGCATCGTGTGAGAACACTACCTTCATGATTCCTCATCCGATGAAGAGGTGAGC 552
 Db 141 GlyPheGlnleuGlnAspPheThrleuylsValAlaIYrilleProAspGlnhrAla 160
 QY 553 -----TCCCTTCGCCCCCTCGAGAGCCAGCGTGGGAC-----CACTTTC 597
 Db 161 GlnGlnAsnProSerProGlnleuArgylArgArgylProGlyGlnArgylSerSer 180
 QY 598 CGGAGCAAGGCGCGCCCTGGGGGCACTTCTCAGGCGAGACAGATTGATTCGCCGCG 657
 Db 181 ArgGlnAla-----SerProGlySerValSerlylGlnlylserProCysAspneuProleu 198
 QY 658 CGGATCCGTCGCCACCCAGTTGTTGGTCATCATCGGAAAGAGGCGTTGACCATTA 717
 Db 199 ArgleuLeuValProThrGlnPheValGlyAlaIlelleGlylsGlnlylAlaThrille 218
 QY 718 AAGAAATCATGTAAGAGAGCCAGTCCCGGATGATATCATGAAAGAAAGAACTCTGGA 777
 Db 219 ArgAsnilleThrlylGlnhrGlnserlylAspValHlssArgylsGlnleuThrlyl 238
 QY 778 GCTGCAAGAGGCTTGTACCATTCATGCCACCCGAGAGGAGCTTCTGAAGATCGCG 837
 Db 239 AlaAlaGlnlylserlleThrilleuSerThrProGlnlylThrserAlaIaCyslys 258
 QY 838 ATGATCTTGAATCATGATGAGAGAGGAGGATGAGCAATAGCCGAGAGATTCCT 897
 Db 259 SerlleuGlnlyllethetHlssylsGlnAlaGlnAspIlelyshethrGlnlylPro 278
 QY 898 CTGAAATCTGGACACATGAGCTGTTGAGAGACTGATGGAAGAAAGAGCAAT 957
 Db 279 LeuyllelleuAlaHlssAsnAsnPhelValGlyArgleuIlelleGlylsGlnlylArgasn 298

```
QY 958 TTGAAGAAATTTGAACATGAAACAGGACCAAGATTAACATCTCATCTTTCGAGATTGG 1017
DB 299 LeuLysLysIleGlnGlnAspThrAspThrLysIleThrIleSerProLeuGlnGlnLeu 318
QY 1018 AGCATATACACCCGGAAGAACATCACTGTGAAGGCGACAGTTGAGCCTGTGCCAGT 1077
DB 319 ThrLeuTyAspProGlnArgThrIleThrValLysGlnSerValGlnThrCysAlaLys 338
QY 1078 GGTGAGATAGAGATTATGAGAGAGTGCGTGAGCCTTTGAAATATGATGCTGGCTGT 1137
DB 339 AlaGlnGlnGlnIleMetLysLysIleArgGlnSerTyGlnAsnAspIleAlaSerMet 358
QY 1138 AACACCCAC----- 1146
DB 359 AsnLeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPheProPro 378
QY 1147 -----TCCGAGATCTTCCAGCCTGTACCCCATCAGCAG 1182
DB 379 ThrSerGlyMetProProProThrSerGlyProProSerThrLeuThrPro----- 395
QY 1183 TTTGGCCCGTCCCGCATCATCCTCTTATCCAGACGAGGATTTGATCTTCTCATC 1242
DB 396 -----ProTyProGlnPheGlnGln---SerGlnThrGlnThrValHisLeuPheIle 412
QY 1243 CCAACCCAGGCTGTGGGCGCCATCATCGGGAAGAGGCGGCACACATCAACAGCTGGCG 1302
DB 413 ProAlaLeuSerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSer 432
QY 1303 AGATTGCGCGAGCCTCTATCAAGATTGCCCTGCGGAAGCGCAGAGCTCAGCGCAAG 1362
DB 433 ArgPheAlaGlyAlaSerIleLysIleAlaProAlaGlnAlaProAlaLysValArg 452
QY 1363 ATGTCATCATCACCGGCGCCACCGGAAGCCGATTCAGGCCCGAGGACGGATCTTTGGG 1422
DB 453 MetValIleIleThrGlyProProGlnAlaGlnPheLysAlaGlnGlyArgIleTyGly 472
QY 1423 AAACGAAAGAGAGAACTTCTTTAACCCCAAGAGAGAGAGAGAGCTGGAAGCGCATTC 1482
DB 473 LysIleLysGlnGlnAsnPheValSerProLysGlnGlnValLysLeuGlnAlaHisIle 492
QY 1483 AGAGTCCCTCTTCCACAGCTGCGCGGATGGCAAGGTGGCAAGCCGTGACGAA 1542
DB 493 ArgValProSerPheAlaAlaGlyArgValIleGlyLysGlnGlyLysThrValAsnGln 512
QY 1543 CTGCAAGAACTTAACCAAGTGCAGAGTCACTGCTGCTGTCACAAAGCCGAGTGAAT 1602
DB 513 LeuGlnSerLeuSerSerAlaGlnValValAlaProArgAspGlnThrProAspGlnAsn 532
QY 1603 GAGGAAGTATGATCGTACAGATTAATGCGGCACTTCTTTGCTAGCCAGACTGCAAGGCGAG 1662
DB 533 AspGlnValValValLysIleThrGlyHisPheTyAlaCysGlnValAlaGlnArgLys 552
QY 1663 ATCAGGGAATTTGACAAACAGTGAAGCAGCAGCAGAGAA 1704
DB 553 IleGlnGlnIleLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 14
000425 PRELIMINARY; PRT, 579 AA.
AC 000425;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Putative RNA binding protein KOC (KOC).
GN KOC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mueller-Pillasch F., Lacher U., Wallrapp C., Et Al.;
RL Oncogene 0:0-0(0).
```

```
RN [2]
RP SEQUENCE FROM N.A.
RC TISUB=Pancreas;
RA Mueller-Pillasch F., Lacher U., Wallrapp C.;
RL Submitted (0CF-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; U97188; AAC35208.1; -.
DR EMBL; U76705; AAD09223.1; -.
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0003723; F:RNA binding; TAS.
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
DR GO; GO:0006412; P:protein biosynthesis; TAS.
DR GO; GO:0006396; P:RNA processing; TAS.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR005054; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
DR PROSITE; PS00300; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 579 AA; 63720 MW; AE5C3A8EE3C135C5 CRC64;
```

Alignment Scores:

Pred. No.:	1,14e-128	Length:	579
Score:	1868.00	Matches:	368
Percent Similarity:	78.52%	Conservative:	78
Best Local Similarity:	64.79%	Mismatches:	96
Query Match:	31.74%	Indels:	26
DB:	4	Gaps:	6

US-09-270-437D-8 (1-3283) x 000425 (1-579)

```
QY 73 ATGAACAAGCTTTTCATCGGGAACCTGAGCCCGCCGTCACCGCGAGACTCCGCGAG 132
DB 1 MetAsnLysLeuTyIleGlyAsnLeuSerGlnAsnAlaAlaProSerAspLeuGlnSer 20
QY 133 CTCTTTGGGACAGAGAACTGCCCCGTGGGACAGAGCTCTGTAAGTCCGCGTACGCC 192
DB 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyAla 40
QY 193 TTCGTGACCTACCCCGACAGAACTGGCGCATTCGCGCATCGAACCTCTCGGGTAA 252
DB 41 PheValAspCysProAspGlnSerThrAlaLeuLysAlaIleGlnAlaLeuSerGlyLys 60
QY 253 GTGGAATTGCAATGGGAAATATCATGAGTTGATTACTCACTCTTAATAAAGCTAAGAGC 312
DB 61 IleGlnLeuHisGlyLysProIleGlnValGlnHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCACTGAGTGGAGAGTGTGATGATGACTT 372
DB 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGlnValLeuAspSerLeu 100
QY 373 TTGCTCAATATGGGACAAGTGAAGATGTGAACAGTCAACAGACACACAGAAACCGCC 432
DB 101 LeuValGlnTyArgIleValAlaGlnSerCysLeuGlnValAsnThrAspSerGlnThrAla 120
QY 433 GTTGTCACAGCTCACTATGCAACAAGAGAAACAAATAAGCCATGAGAACTAAGC 492
DB 121 ValValAsnValThrTySerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAAGACTACTCTTCAAGATTTCCTACATCCCGGATGAAGAGGTGAGC 552
DB 141 GlyPheGlnLeuGlnAsnPheThrLeuLysValAlaIleTyIleProAspGlnMetAlaIle 160
QY 553 TCCCTTCGCCCTCAG-----CGAGCCAGCGT-----GGGAGCAACTCTTCCGG 600
DB 161 GlnGlnAsnProLeuGlnGlnProArgIleArgArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAAGGCCACCGCCCTGGGGGAGCACTTTCAGAGCCAGACAGATTTGATTCGCGTGGG 660
```


QY 643 ATGATGTTCCCGGCGGAGATCCGTCGTCGCCACCACTGTTTGTGTGCATCATCGAAG 702
 Db 83 AlaSppheProLeuArgIleLeuValGlnSerGluMetValGlyAlaIleIleGlyArg 102
 QY 703 GAGGCGCTTGACATTAAGAAACATCACTAAGACACACCGAGCCGGGTAGATACCTA 762
 Db 103 GlnGlySerThrIleArgThrIleThrGlnGlnSerArgAlaArgValAspValHisArg 122
 QY 763 AAGAGAACTCTGGAGCTCGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 822
 Db 123 LysGluAsnValGlySerLeuGlnLysSerIleThrIleTyrGlyAsnProGluAsnGly 142
 QY 823 TCTGAAGCATGGCCGACATGATTTCTTGAATATCATGACAAAGAGGACAGATAGACCA 882
 Db 143 ThrAsnAlaCysLysArgIleLeuGlnValMetGlnGlnGlnAlaIleSerThrAsnLys 162
 QY 883 GCCGAA-----GAGATTCCTCGAAATCTGGCCACACATGAGCTTG 924
 Db 163 GlyGlnLeuSerProGluCysSerGlnIleCysLeuLysIleLeuAlaHisAsnAsnLeu 182
 QY 925 GTTGGAAAGCTGATTTGAAAAGAGGACAGAAATTTGAAGAAATTTGAACATGAACAGG 984
 Db 183 IleGlyArgIleIleGlyLysSerGlyAsnThrIleLysArgIleMetGlnAspThrSer 202
 QY 985 ACCAAGATATACAATCTCATCTTTGACAGATTGAGCATATACACCCGGAGAAACCATC 1044
 Db 203 ThrLysIleThrValSerSerIleAsnAspIleAsnSerPheAsnLeuGlnLysIleIle 222
 QY 1045 ACTGTAAGAGGACAGTGTGAGCGCTGTGCGACATGCTGATATGAGATTTATGAAGAGCT 1104
 Db 223 ThrValLysGlyLeuIleGlnLysMetSerIleArgAlaGlnAsnGlnIleSerThrLysLeu 242
 QY 1105 CGTGAAGCCTTTGAATAATGATATGCTGCGTGTTTAACACCCACTCCGATACTCTTCACG 1164
 Db 243 ArgGlnSerTyrGluAsnAspLeuGlnAlaIleAlaProGlnSerLeuMetPheProGly 262
 QY 1165 CTGTACCCCATCACCACTTTGGC----- 1188
 Db 263 LeuHisProMetAlaMetMetSerThrProGlyAsnGlyMetValPheAsnThrSerMet 282
 QY 1189 CCGTTCGGCGCATCATCACTCT----- 1209
 Db 283 ProPheProSerCysGlnSerPheAlaMetSerLysThrProAlaSerValValProPro 302
 QY 1210 ---TATCCAGAG-----CAGAGATTTGTGATCTCTTCATGCCAACCCAGAGCTGTGGCG 1260
 Db 303 ValPheProAsnAspLeuGlnGlnIleThrThrTyrLeuTyrIleProAsnAsnAlaValGly 322
 QY 1261 GCCATCATCGGGAGAGAGAGGGGGACACATCAACACAGCTGGGAGATTGCGCGAGCTCT 1320
 Db 323 AlaIleIleGlyThrArgGlySerHisIleArgSerIleMetArgPheSerAsnAlaSer 342
 QY 1321 ATCAAGATTGGCCCT-----GGGAGAGGCCA-----GACGTACAGGAAAGATGGTC 1368
 Db 343 LeuLysIleAlaProLeuAspAlaAspLysProLeuAspGlnGlnIleThrGlnArgLysVal 362
 QY 1369 ATCATCAACCGGACACCGGAAAGCCAGATTCAGAGCCAGGACCGATCTTTGGAAATCG 1428
 Db 363 ThrIleValGlyThrProGlnGlyGlnThrLysAlaGlnTyrMetIlePheGlnLysMet 382
 QY 1429 AAAAGAGAAAATTCTTTAACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1488
 Db 383 ArgGlnGlnGlyPheMetCysGlyThrAspAspValArgLeuThrValGlnLeuLeuVal 402
 QY 1489 CCCCTCTCAACAGCTGGCGGGGTATTTGGCAAAAGTGGACACCGTGAAAGAACTGGAG 1548
 Db 403 AlaSerSerGlnValGlyArgIleIleGlyLysGlyGlnAsnValArgGlnLeuGln 422
 QY 1549 AACTTAAACAGAGAGAGATCATGCTGCTGCTGACAA---ACGGCAAGATGAA----- 1599
 Db 423 ArgValThrGlySerValIleLysLeuProGlnHisAlaLeuAlaProProSerGlyGly 442
 QY 1600 AATAGAGAAAGTATGTCGCAAGATTAATGAGGCACTTTTGTGTAGCCAGACTGACAGAGCG 1659

[illegible]

Search completed: July 23, 2004, 11:15:36
Job time : 192.961 secs

This Page Blank (uspto)